



The Alberta Journal of Educational Research

Volume XXI, No. 4

December, 1975

CONTENTS

Toward Teacher Education in the Year 2000	221
<i>S. C. T. Clarke and H. T. Coutts</i>	
Classroom Climate and Achievement in Secondary School Mathematics Classes	241
<i>R. O'Reilly</i>	
Classroom Personality Patterns of Teachers of Secondary Mathematics	249
<i>D. F. Robitaille</i>	
Socratic Dialogue vs. Behavioural Practice in the Development of Coping Skills	255
<i>J. P. Klein</i>	
Methodological Problems in Conservation Testing with Particular Reference to Volume Conservation	262
<i>E. D. Hobbs</i>	
An Investigation Into the Relationship of the Pausing Phenomena in Oral Reading and Reading Comprehension	278
<i>Sr. Ruth Eagan</i>	
Rural Standardisation of the Burt-Vernon Graded Word Reading Test	289
<i>G. Broadley and K. M. Broadley</i>	
Teacher Self-Actualization and Pupil Control Ideology	295
<i>L. E. Jury, D. J. Willower and W. J. DeLacy</i>	

FACULTY OF EDUCATION
The University of Alberta

S . C . T . C L A R K E

and

H . T . C O U T T S

The University of Alberta

Toward Teacher Education in the Year 2000

A modified Delphi technique and Q-sort was used to replicate and update a previous study conducted four years earlier. The views of chief administrative officers in English-language teacher education institutions in Canada about the future of teacher education were sought. The 222 statements obtained from the participants were classified and 36 representative statements about the future of teacher education were educed. The impact on education of the feature or aspects described in the statements were obtained from the Q-sort.

A feature or aspect of teacher education in the future originated by many of the participants and rated as having most impact on education was that it will be non-terminal, continuing, and lifelong. The feature judged to have next most impact on education was that teacher education in the future would feature a closer relationship between theory and practice. The study presents additional features and ratings.

There were 41 participants representing 32 of the 39 English-language teacher education institutions in Canada. (Dr. Clarke is Director of Special Sessions and Dr. Coutts is Professor Emeritus at The University of Alberta.)

Overview

The study reported here was designed in part to replicate, in part to update in 1974-75, a previous one conducted by the authors (1970-71).

In the present study a modified Delphi technique was used to elicit from the chief administrative officers of the English-language teacher education institutions in Canada their views about features or aspects of teacher education in the year 2000. The 222 statements obtained from the participants were classified, and 36 were educed to represent these features or aspects. Participants then rated these features or aspects as to probable date of acceptance, desirability, and feasibility, and were encouraged to comment on each. As a third activity, participants completed a seven-place Q-sort in which they expressed opinions concerning the impact on education of each of the 36 stated features or aspects of teacher education.

The most frequently mentioned features or aspects originated by the 45 participants were: multiple programs, lifelong or continuing involvement, humanistic concerns, increased emphasis on the school setting, a stronger or more scholarly data base, and increased selectivity. About ten percent of the comments which participants made on the 36 features or aspects of teacher education complained of ambiguity because of the use of terms such as "broader data base" and "humanistic concerns." Despite these difficulties in communication, 41 out of 45 participants completed all three parts of the study.

The five features or aspects of teacher education toward the year 2000 judged to have most impact on education, their ratings, the comments made by participants, and the number of participants who originated statements of this kind, follow:

"A recognized and accepted aspect of teacher education in the future will be that it is non-terminal, continuing, lifelong." This feature or aspect was rated as having most impact on education. Twenty-two percent of participants said it is in effect now, 22% predicted it will be before 1980, and 51% predicted it will be in effect in the majority of teacher education institutions in Canada between 1980 and 2000. Participants saw this as desirable, but moderately difficult to accomplish. Forty-two percent of the participants originated statements stating or implying this feature or aspect.

"Teacher education in the future will feature a closer relationship between theory and practice." Second in its impact on education, this feature or aspect of teacher education was judged by the participants to be in effect now or to be likely to be so before 1980. Participants saw this as very desirable, but difficult. Twelve percent of participants originated a statement along the lines of this feature or aspect.

"In the future continuing (inservice) teacher education will be equal in importance to preservice teacher education." This feature or aspect was rated as having third most impact on education. Participants were divided as to when this might be in effect in the majority of teacher education institutions in Canada. Twenty-seven percent said before 1980, 47% said between 1980 and 2000, 13% said after 2000, and 9% said never. This feature or aspect of teacher education was seen as very desirable. Participants were divided about feasibility: 22% said it was easy or very easy, 29% rated it "neutral," and 45% said it was difficult or very difficult. Forty-two percent of participants originated a statement about the importance of continuing education in the future, but not necessarily with the incisiveness of this one.

"Teacher education in the future will feature a large number of differentiated programs designed to fit the capabilities, previous experience, and career aspirations of candidates as well as their subsequent functions in newly-developed school programs." While rating this feature or aspect fourth in its impact on education, a few participants said it was in effect now, 33% said it would be before 1980, and 47% predicted it would be in a majority of the teacher education institutions in Canada between 1980 and 2000. It was rated as very desirable, but difficult. Forty-two percent of participants originated statements along the lines of this feature or aspect.

"Teacher education in the future will provide for a more rigorous screening of candidates on the basis of evaluative criteria which take into account personal as well as academic qualifications." Rating this feature or aspect fifth in its impact on education, the majority of participants saw this as in effect now or by 1980. Although 27% thought it would not be in effect in a majority of the teacher education institutions in Canada before 1980-2000, another 9% thought after 2000, and another 7% never. It was rated as very desirable and very difficult. Thirty-one percent of participants originated statements dealing with selection.

Comparisons with the previous study (1970-71) revealed an overlap of only four (about 10%) participants. The most frequently mentioned features or aspects of teacher education in the first study were also mentioned in the second. There was less correspondence in the two studies of features or aspects that had been originated by only a few participants.

Comparisons with studies of the goals of education reveal that there is no one-to-one correspondence between the goals of education as depicted in various studies and the preparatory programs of the teachers who implement those goals. The authors believe that both goals and features or aspects of the future of teacher education express a *felt need*; that is, participants do not list the most assured or likely future events nor specify as a goal that which is taken for granted as occurring.

A comparison between one study on the goals of teacher education (Coutts & Clarke, 1972) and the present study of teacher education toward the year 2000 showed a high degree of congruence.

Population

The chief administrator in each English-language teacher education institution in Canada was invited to participate, and in the event that the institution enrolled more than 1,000 students, to name an additional person for each 1,000 students or fraction thereof.

Of the chief administrators of the 39 institutions contacted, 32 agreed to participate. Of the 20 nominees of the chief administrators who agreed to participate, 4 were assistants, associates, deputies, or equivalent; 14 were department chairmen or heads, directors, coordinators, or equivalent; and 2 had no administrative title. Seven institutions originally contacted were not represented in the study.

The views of teacher education toward the year 2000 as expressed in this study are the views of the chief administrators and their immediate colleagues in 32 out of 39 of the English-language teacher education institutions in Canada. The 32 institutions included 91% of the students enrolled.

Statements about the future of teacher education

Each of the 52 persons who agreed to participate was sent an offprint of the *Journal of Teacher Education* article which described the 1970-71 Clarke-Coutts study and a single sheet with these instructions: "Thinking about the future as we approach the 21st century, what will teacher education be like? You may wish to precede each statement you make with the stem 'I believe that teacher education in the future will . . .'" The sheet bore the title of this study and a subtitle "STATEMENTS ABOUT

TEACHER EDUCATION IN THE FUTURE." Most of the participants (89%) produced 4, 5, or 6 statements. The range was 3 to 10.

The 222 statements obtained were classified, as shown in Table 1, in order to facilitate the educing of 36 representative items. Since the basic purpose of the classification was to facilitate the educing and phrasing of representative statements, no special efforts were made to refine or validate the classes. In some instances, there was little difficulty about classification, e.g., that in the future there would be more emphasis on selection: initially, throughout preservice teacher education, at the end before placement, etc. In other instances, the classification was difficult, e.g., as between humanistic concerns and stress on personal development. Some statements made by participants contained more than one idea. It was decided that the core or central idea, as judged by the authors, should be used as a single statement because this would seem to be a more reasonable representation of the views of the different participants than would result from breaking up omnibus statements into two, five, or possibly ten ideas included. This decision made classification more difficult.

For the reasons stated above undue importance should not be placed on the classes and the number of statements presented in Table 1. It is the opinion of the authors, however, that the classification is sufficiently valid to warrant the presentation of the classes and the frequency of statements therein as a first approximation to the future of teacher education as seen by the participants. The reader is reminded again that these participants were the chief administrators in most of the English-language teacher education institutions in Canada.

There were 45 participants who submitted statements. The commonality of statements was not great: at most 19 out of 45 participants (or 42%) originated a statement of the same classification. Six classes included 105 (53%) of the 199 statements classified as other than miscellaneous. Viewed in this way, one can detect a preoccupation with multiple programs, lifelong teacher education, humanistic concerns, school setting, a stronger data base, and selection, in the statements which came to the minds of the chief administrators as they contemplated teacher education about the year 2000.

Representative statements

The task of educing a set of statements about teacher education in the future representative of those originated by the participants drew heavily on the scholarship, impartiality, and command of English usage of the authors. The classification, made for this purpose, simplified the task. Admittedly statements should be clear and unambiguous, should each present one and only one idea, and should be representative of the original set. Because participants would be asked to rate and comment on the statements in Part II and to sort them in Part III of the study, it was decided arbitrarily to limit the number to 36. The statements in each classification were examined and the number of items necessary to represent the ideas within various categories was determined; for example, the first three statements are based on the "Multiple Programs" classification. It should be emphasized that the procedure used focussed on majority views;

TABLE 1
CLASSIFICATION OF STATEMENTS ABOUT THE FUTURE OF
TEACHER EDUCATION

Classification	Number of Statements
1. MULTIPLE PROGRAMS: individualized (8), differentiated but not individualized (11)	19
2. LIFELONG OR CONTINUING: using the term "inservice" (10), using the term "continuing" (9)	19
3. HUMANISTIC: human relations, interpersonal skills, communication skills, commitment to social needs	19
4. INCREASED EMPHASIS ON SCHOOL SETTING: internship (10), field experience (6), apprenticeship (2)	18
5. STRONGER OR MORE SCHOLARLY DATA BASE: more knowledge about human behaviour	16
6. MORE SELECTIVE	14
7. ONE OUTCOME STRESSED: diagnostician, remediator, curriculum developer, tutor, team member, facilitator of learning, etc.	11
8. MORE SPECIALIZATION	9
9. CONTROL AND RESPONSIBILITY INCREASINGLY SHARED: with teachers' organizations, school boards, government, etc.	9
10. CONTROL AND RESPONSIBILITY TO MOVE OUT OF THE HANDS OF THE UNIVERSITY	7
11. COMPETENCY BASED EMPHASIS	7
12. TEACHERS TO BE PREPARED FOR ROLE ADDITIONAL TO TEACHING GRADES 1-12 (13)	7
13. INCREASED EMPHASIS ON THE COMMUNITY SETTING	7
14. INCREASED TOTAL-UNIVERSITY INVOLVEMENT	6
15. GREATER RESEARCH ORIENTATION	6
16. INTERDISCIPLINARY EMPHASIS, INTEGRATED, LESS FRAGMENTED	5
17. INCREASED EMPHASIS ON PERSONAL DEVELOPMENT	5
18. UNIFIED PROGRAMS: marriage of theory and practice	5
19. MORE EMPHASIS ON PREPARING GENERALISTS	4
20. GREATER EMPHASIS ON VALUES	3
21. LONGER	3
22. MISCELLANEOUS: not classifiable in above categories, largely because they were statements about education	23
TOTAL	222

highly individual views could be lost within a classification or because of the 36 statement limitation.

The educed statements were listed in Part II of the study in order as derived from Table 1 with those representing frequent mention coming first, and those representing infrequent mention coming last. Two items which kept reappearing as asides in statements which were classified elsewhere were included (numbers 33 and 34): teacher education in the future will be about the same as it is now, and teacher education in the future will bear little resemblance to what it is now.

Comments on statements

Part II invited participants, as they rated the items, "to comment on your decision in the space provided." Only four of the participants made no comments. At the other extreme, six participants commented on every item. The average number of comments was 16.5.

There were 73 comments classified as Communication, or an average of 1.7 per participant; 129 classified as Explanation of Rating, or an average of 2.9 per participant; and 541 classified as Comments on Substance for an average of 12.0 per participant. The difficulty in communication is illustrated in item 10, "Teacher education in the future will feature a much greater emphasis on humanistic concerns," where 10 participants (22%) found the term *humanistic concerns* too vague. An examination of the 36 items reveals some where the expectation of possible ambiguity is high, but where none was mentioned; for example, item 22, "Teacher education in the future will involve the community and community resources to a much greater extent." Insofar as participants experienced difficulty in communication, their willingness to rate the items was reduced.

The comments which explained the participant's ratings are self-explanatory. Those classified as substantive were observations about the cost, resources required, social pressures, agreed-upon goals, and trends which the item generated in the minds of the participants.

The comments reveal that lack of clarity of meaning casts doubt on the finding associated with items 10, 12, and 31. Yet such is the idiosyncratic nature of communication that many participants appeared to experience no difficulty with these items. One of the most interesting findings is the small variation in the number of comments per item. A careful perusal of the comments does reveal an individual interpretation on many of the items which may mean that there is less true agreement than the findings show.

Ratings

In Part II of the study each participant was asked to rate the probable date of acceptance, desirability, and feasibility of each item. Table 2 present the items, the results in numbers and a descriptive analysis. For the first item, "The basic program of teacher education in the future will vary in length for different candidates," 4 participants (9%) indicated that this is in effect now, 15 (33%) that it will be before 1980, 13 (29%) that it will happen between 1980 and 2000, 8 (18%) that it will occur after 2000, and 3 (5%) that this statement about the future of teacher education would never hold. There were two (4%) participants who did not rate the probable date of acceptance, for a total of 45 participants.

The analysis of the comments in the preceding section revealed that for many of the participants items 10, 12, and 31 were not clear. An inspection of the "not rated" column of Table 2 shows that for item 10, 8 participants (18%) did not rate desirability or feasibility, and 9 participants (20%) did not rate probable date of acceptance. For item 12, the numbers who did not rate the corresponding parts were 15, 15, and 13 respectively, and for item 31 were 8, 7, and 7 respectively. Obviously, one reaction of participants was to decline to rate a statement which they did not understand. Two other items (33 and 34) frequently were not rated because they did not fit the

rating system, in that one stated that teacher education in the future would be about the same as it is now, while the other stated that it would bear little resemblance to what exists in 1975. Three other items were not rated by a number of participants: 18 and 19 which dealt with who, in the future, would control teacher education, and 26 which predicted emphasis on preparation of generalists.

Using a simple weighting where very desirable is assigned a weighting of 5, desirable 4, neutral 3, undesirable 2, and very undesirable 1, the features or aspects of teacher education found most desirable were as follows in rank order of the combined weightings for "very desirable" and "desirable": #7 (lifelong or continuing involvement), #27 (closer relationship between theory and practice), #9 (emphasis on human relations), #3 (differentiated programs), #29 (interdisciplinary programs and procedures) and #6 (greater proportion to resources to inservice) tied, #15 (stress on personal development), #16 (total university commitment), #22 (involvement of the community), and #2 (individualization).

Using a similar weighting system with respect to difficulty, rank order in decreasing difficulty is as follows: #23 (teaching practice linked to apprenticeship of health service neophytes), #14 (selection), #16 (total university commitment), #3 (differentiated programs), #15 (stress on personal development), #27 (closer relationship between theory and practice), #28 (knowledge focussed on applied skills), #13 (emphasis on research) and #29 (interdisciplinary programs and practices) tied, and #2 (individualization).

There is a heavy overlap between the ten most desirable and the ten most difficult items, #27 (closer relationship between theory and practice), #3 (differentiated programs), #29 (interdisciplinary programs and practices), #2 (individualization), #15 (stress on personal development), and #16 (total university commitment) being on both lists.

At the other end of the scale, the five features or aspects judged as least desirable were: #33 (teacher education will remain as it is now), #19 (move out of the hands of the universities), #17 (joint control by teacher organizations and teacher education) and #18 (joint control: universities, professional organizations, employing agencies, government) tied, and #20 (centre on the development of competencies).

The five rated as most easily accomplished were: #9 (emphasis on human relationships), #16 (a greater proportion of resources to inservice), #1 (varying length of program), #4 (more time in schools), and #22 (community involvement).

We can now identify two items deemed to be desirable (in the top ten) and easy (in the top 5): #9 (emphasis on human relations) and #6 (a greater proportion of resources to inservice).

The rating of probable date of acceptance permitted participants to indicate disagreement with a feature or aspect which had been originated by other participants. In descending order, the following features or aspects were rated as "never" for probable date of acceptance: #19 (teacher education will move out of the hands of the universities), #33 (teacher education will remain as it is now), #18 (joint control: universities, professional organizations, employing agencies, government), #17 (joint control by teacher organizations and teacher education institutions), #30 (longer), #25

Table 2
Ratings of Probable Date of Acceptance, Desirability, Feasibility, and of Impact on Education
of Thirty-six Features or Aspects of Teacher Education toward the Year 2,000

Features or Aspects of Teacher Education toward the Year 2,000	Probable Date of Occurrence				Desirability				Feasibility					Not Rated		Impact on Education							Wtd Total	Rank Order			
N = Number Wtd = Weighted																Q-Sort Frequency MOST....to....LEAST											
																Weighting											
	Now	Before 1980	1980-2000	After 2000	Never	Very Desirable	Desirable	Neutral	Undesirable	Very Undesirable	Very Easy	Easy	Neutral	Difficult	Very Difficult	Date	Desirability	Feasibility	7	6	5	4			3	2	1
1. The basic program of teacher education in the future will vary in length for different candidates.	N %	4 9	15 33	13 29	8 18	3 7	18 40	17 38	7 16	2 4	0 0	3 7	16 36	6 13	14 31	5 11	2 4	1 2	1 4	1 2	3 16	4 13	3 4	2 3	1 3	142	31
2. Teacher education in the future will be much more individualized, with candidates having much more say in the design of their own programs.	N %	2 4	13 29	28 62	1 2	0 0	18 40	23 51	2 4	1 2	0 0	1 2	12 27	5 11	26 58	0 0	1 2	1 2	1 2	2 2	2 12	2 21	2 3	3 0	0 175	11	
3. Teacher education in the future will feature a large number of differentiated programs designed to fit the capabilities, previous experience, and career aspirations of candidates as well as their subsequent functions in newly-developed school programs.	N %	2 4	15 33	21 47	5 11	2 4	21 47	21 47	1 2	0 0	1 2	7 16	4 9	27 60	4 9	4 0	1 2	2 4	0 2	1 2	5 20	5 5	0 0	0 0	0 191	4	
4. In the future candidates will spend a notably greater portion of the time devoted to their programs in schools.	N %	9 20	25 56	11 24	0 0	0 0	20 44	17 38	5 11	3 7	0 0	2 4	17 38	14 31	11 24	1 2	0 0	0 0	0 0	1 4	12 15	7 7	2 2	0 0	0 176	9	
5. Teacher education in the future will include a supervised internship with a total duration of approximately a year.	N %	5 11	12 27	21 47	3 7	1 2	20 44	15 33	7 16	2 4	0 0	2 4	6 13	15 33	19 42	1 2	3 7	1 2	2 4	4 4	12 14	9 9	2 2	0 0	0 169	19	
6. Teacher education in the future will devote a greater proportion of its total resources to inservice education.	N %	7 16	19 42	15 33	1 2	1 2	27 60	13 29	3 7	1 2	0 0	1 2	19 42	10 22	12 27	2 4	2 4	1 2	1 2	1 5	16 15	2 2	2 2	0 0	0 187	6	

Features or Aspects of Teacher Education toward the Year 2,000		Probable Date of Occurrence				Desirability				Feasibility				Not Rated			Impact on Education							Wtd Total	Rank Order		
		Now	Before 1980	1980-2000	After 2000	Never	Very Desirable	Desirable	Neutral	Undesirable	Very Undesirable	Very Easy	Easy	Neutral	Difficult	Very Difficult	Date	Desirability	Feasibility	7	6	5	4			3	2
N = Number Wtd = Weighted																											
7. A recognized and accepted aspect of teacher education in the future will be that it is non-terminal, continuing, lifelong.	N %	10 23	10 22	23 51	2 4	0 0	37 82	7 16	0 0	1 2	0 4	12 27	15 33	5 11	5 11	0 0	0 2	1 2	3 9	11 14	3 3	1 1	0 0	197	1		
8. In the future continuing (inservice) teacher education will be equal in importance to preservice teacher education.	N %	1 2	12 27	21 47	6 13	4 9	24 53	11 24	6 13	2 4	0 0	1 2	9 20	13 29	7 16	1 2	2 4	2 7	13 16	2 2	1 1	0 0	193	3			
9. In the future there will be an increased emphasis on human relations in preparing candidates for teaching.	N %	10 22	19 42	15 33	0 0	1 2	28 62	14 31	3 7	0 0	0 0	3 7	18 40	6 13	13 29	5 11	0 0	0 0	2 8	7 16	6 6	2 2	0 0	183	8		
10. Teacher education in the future will feature a much greater emphasis on humanistic concerns.	N %	4 9	14 31	15 33	0 0	3 7	21 47	13 29	3 7	0 0	0 0	2 4	11 24	8 18	5 11	9 20	8 18	3 1	7 23	5 5	2 2	0 0	173	15			
11. Teacher education in the future will be based on a hard central core of knowledge (nature of development, learning theory, communication and social interaction theory, educational design, etc.) and professional skills, with a corresponding diminution of "conventional wisdom."	N %	3 7	10 22	15 33	10 22	5 11	17 38	14 31	7 16	5 11	0 0	1 2	9 20	12 27	8 18	2 4	2 4	1 7	8 10	13 13	2 2	0 0	174	13			
12. Teacher education in the future will rest on a broader data base and will demand that candidates possess a mastery of this base.	N %	1 2	9 20	13 29	4 9	5 11	7 16	17 38	5 11	1 2	0 0	0 11	13 29	11 24	1 2	13 29	15 33	1 3	4 14	15 15	4 4	0 0	154	25			

Features or Aspects of Teacher Education toward the Year 2,000	Probable Date of Occurrence				Desirability				Feasibility				Not Rated			Impact on Education							Wtd Total	Rank Order																																			
N = Number Wtd = Weighted	Now				Very Desirable				Desirable				Neutral				Undesirable				Very Undesirable				Very Difficult				Date				Desirability				Feasibility																						
	Before 1980				1980-2000				After 2000				Never																																														
13. Teacher education in the future will feature increased emphasis on research in teaching and teacher education.	N 3	13	18	8	1	19	17	7	0	0	1	5	10	23	4	2	2	2	2	2	4	9	17	7	2	0	176	10																															
14. Teacher education in the future will provide for a more rigorous screening of candidates on the basis of evaluative criteria which take into account personal as well as academic qualifications.	N 5	18	12	4	4	30	6	3	2	2	1	4	2	15	21	2	2	2	2	4	5	11	15	6	0	0	191	5																															
15. Teacher education in the future will feature much greater concern for, and devote much more time to, the personal development of candidates.	N 2	15	20	5	1	22	19	1	1	0	1	5	7	23	7	2	2	2	2	1	1	12	19	4	3	1	168	21																															
16. Teacher education in the future will involve much more concern and commitment by the total university.	N 2	6	20	10	5	26	13	4	1	0	1	2	5	19	16	2	1	2	0	2	6	15	10	7	1	147	28																																
17. Teacher education in the future will be under the joint control of teachers' organizations and teacher education institutions.	N 2	2	20	4	14	5	11	14	7	8	1	5	15	11	10	3	0	3	1	2	0	11	12	11	4	125	33																																
18. Teacher education in the future will be controlled jointly by universities, professional organizations, employing agencies, and government, in that order.	N 7	3	6	6	17	6	10	9	7	8	3	6	8	8	13	6	5	7	0	0	3	13	16	9	0	123	35																																
19. In the future, control of teacher education will move out of the hands of the universities.	N 2	5	3	6	21	0	4	12	10	15	2	6	14	5	10	8	4	8	1	0	3	11	9	14	3	124	34																																

Features or Aspects of Teacher Education toward the Year 2,000		Probable Date of Occurrence				Desirability				Feasibility				Not Rated			Impact on Education							Wtd Total	Rank Order			
																	Q-Sort Frequency											
																	MOST....to....LEAST											
																	Weighting											
		Now	Before 1980	1980-2000	After 2000	Never	Very Desirable	Desirable	Neutral	Undesirable	Very Undesirable	Very Easy	Easy	Neutral	Difficult	Very Difficult	Date	Desirability	Feasibility	7	6	5	4	3	2	1		
20. Teacher education in the future will center on the development of competencies required for the successful performance of teaching functions rather than the acquisition of subject matter.	N %	8 18	14 31	8 18	3 7	7 16	11 24	12 27	7 16	7 16	4 9	2 4	13 29	6 13	16 36	3 7	5 11	4 9	5 11	0	5	10	19	5	1	1	174	14
21. Whereas at the present time most candidates are prepared for teaching in Grades 1-12 (or 13), in the future a greater proportion will be prepared for other functions and/or positions.	N %	2 4	12 27	19 42	2 4	5 11	9 20	22 49	4 9	3 7	3 7	0 36	16 36	16 36	7 16	1 2	5 11	4 9	5 11	0	3	4	17	11	5	1	150	26
22. Teacher education in the future will involve the community and community resources to a much greater extent.	N %	4 9	20 44	15 33	2 4	2 4	18 40	19 42	5 11	0 0	0 0	3 7	15 33	9 20	13 29	2 4	2 4	3 7	2 7	1	3	10	19	8	0	0	175	12
23. Teacher education in the future will feature a broad and strong community context with teaching practice linked to the "apprenticeship" of counterparts in the helping professions such as social work, public health, and recreation.	N %	1 2	7 16	24 53	10 22	3 7	7 16	26 58	9 20	1 2	2 4	0 11	5 13	13 29	24 53	3 7	0 0	0 0	0 0	0	2	13	19	4	3	0	171	17
24. Teacher education in the future will feature a greater emphasis than is currently in evidence on value systems and the valuing process.	N %	2 4	18 40	14 31	6 13	2 4	13 29	22 49	9 20	0 0	0 0	0 18	8 31	14 29	13 29	9 20	3 7	1 2	1 2	0	5	22	5	4	0	166	22	
25. Teacher education in the future will feature much more specialized preparation for limited and clearly defined functions and responsibilities.	N %	1 2	11 24	13 29	6 13	12 27	5 11	14 31	12 27	5 11	6 13	1 2	9 20	17 38	7 16	6 13	2 4	3 7	5 11	0	0	6	15	18	1	1	147	29

Features or Aspects of Teacher Education toward the Year 2,000	Probable Date of Occurrence				Desirability				Feasibility				Not Rated			Impact on Education							Wtd Total	Rank Order				
N = Number Wtd = Weighted																Q-Sort Frequency MOST...to....LEAST												
																Weighting												
26. Teacher education in the future will emphasize the preparation of general- ists more than specialists.	N	8	9	2	12	7	18	8	3	4	4	14	13	4	3	6	5	7	0	3	8	11	11	7	1	150	27	
	%	18	18	20	4	27	16	40	18	7	9	9	11	29	9	7	13	11	16						0	195	2	
27. Teacher education in the future will feature a closer relationship between theory and practice.	N	5	21	15	2	1	31	12	1	0	0	3	8	5	17	10	1	2	4	5	10	21	1	0	0	195	2	
	%	11	47	33	4	2	69	27	2	0	0	7	18	11	38	22	2	2	4									
28. Teacher education in the future will feature an integration of knowledge from whatever source to focus on applied skills and competencies req- uired by the effective teacher, rather than a curriculum based on tradit- ional disciplines and courses.	N	3	13	13	8	3	18	15	3	3	2	0	5	9	19	8	5	4	4	1	4	7	21	7	1	0	173	16
29. Teacher education in the future will be increasingly interdisciplinary with contributions from a wide range of specialists in appropriate dis- ciplines being focussed on problems specific to teaching.	N	0	13	23	7	0	19	23	1	0	0	4	12	18	8	2	2	3	1	3	13	17	7	0	0	169	20	
30. Teacher education in the future, while enabling candidates to teach after the completion of a basic pro- gram of teacher education at degree level, will increasingly involve 5 to 8 years of initial preparation.	N	1	4	13	11	13	4	16	9	10	3	3	7	9	15	5	3	3	6	0	2	3	15	16	4	1	144	30
	%	2	9	29	24	29	9	36	20	22	7	7	16	20	33	11	7	7	13									
31. Teacher education in the future will increasingly prepare candidates to be facilitators of learning.	N	9	12	15	2	0	19	17	1	0	0	1	11	4	14	8	7	8	7	2	8	14	8	1	0	184	7	
	%	20	27	33	4	0	42	38	2	0	0	2	24	9	31	18	16	18	16									

Features or Aspects of Teacher Education toward the Year 2,000	Probable Date of Occurrence				Desirability				Feasibility				Not Rated		Impact on Education							Wtd Total	Rank Order				
N = Number Wtd = Weighted	Now	Before 1980	1980-2000	After 2000	Never	Very Desirable	Desirable	Neutral	Undesirable	Very Undesirable	Very Easy	Easy	Neutral	Difficult	Very Difficult	Date	Desirability	Feasibility	Q-Sort Frequency MOST....to....LEAST								
																			Weighting								
																			7	6	5	4	3	2	1		
32. Teacher education in the future will feature the preparation and facilitate the placement of candidates consistent with their personal characteristics.	N %	0 0	6 13	17 38	9 20	8 18	13 29	15 33	6 13	3 7	3 0	12 7	10 27	10 22	15 33	5 11	5 11	5 11	0 1	1 12	19 19	5 5	4 4	0 0	165	23	
33. Teacher education in the future will be just about the same as it was in 1975.	N %	3 7	5 11	1 2	0 42	19 0	0 0	0 0	5 11	17 38	6 13	7 16	8 18	4 9	3 7	17 38	6 13	17 38	1 1	1 0	0 6	4 4	7 7	22 2	85	36	
34. Teacher education in the future will bear little resemblance to teacher education in 1975.	N %	1 2	1 2	8 18	17 38	3 7	7 16	17 38	7 16	4 9	1 2	0 4	2 22	10 20	9 24	11 33	15 20	9 29	1 1	5 13	11 11		8 2	141	32		
35. Teacher education in the future will involve increased use of technology such as computer assisted instruction, instructional packages, videotaped materials, etc.	N %	5 11	18 40	16 36	6 13	0 0	7 16	27 60	9 20	2 4	0 0	1 2	13 29	11 40	18 40	2 4	0 0	0 0	0 0	2 8	14 14	3 3	0 0	156	24		
36. Teacher education in the future will feature an increasing emphasis on controlled clinical experiences (micro-teaching, simulation, tutoring, highly structured and analyzed mini teaching episodes, and the like) in contrast with traditional and/or extended practicums.	N %	5 11	9 20	17 38	2 4	7 10	13 29	16 36	6 13	2 4	5 11	16 36	17 38	1 2	5 11	3 7	4 9		0 4	11 17	6 6	3 3	0 0	171	18		

(more specialized) and #26 (prepare generalists) tied, #32 (preparation and placement consistent with personality), #20 (centre on the development of competencies) and #36 (increased emphasis on clinical experiences). The number of persons who rated the feature or aspect of teacher education as never likely to occur decreased from 21 who so rated item 19 to 7 who so rated item 36.

As might be expected, features or aspects of teacher education rated as least desirable are also rated as never likely to be accepted or occur. All five of the least desirable items (33, 19, 17, 18, and 20) are found in the ten rated least likely to be accepted or occur.

Further study of Table 2 reveals that there was, indeed, a definite future cast in the 36 statements. Very few participants indicated that the feature or aspect of teacher education described was in effect now. For two, #7 (lifelong or continuing) and #9 (emphasis on human relations) 10 participants (22%) indicated that this feature was in effect now. All other items had a lower ratings.

The analysis of participant's ratings in Table 2 summarized the major reactions to the features or aspects of teacher education delineated by the 36 items according to the preceding analysis.

Impact on education

In Part III of the study, a specially prepared Q-sort requested participants to select the one feature or aspect of teacher education with most impact on education, the three with next most impact, then the seven, fourteen, seven, three, and finally the one with least impact on education. Forty-one of the 45 participants who had completed Parts I and II completed Part III.

Participants observed that it was difficult to include two statements "that teacher education in the future would be just about the same as now" or "will bear little resemblance to teacher education in 1975." The items previously noted as not being clear might have caused difficulty in the Q-sort, but a perusal of Table 2 shows that many participants were able to sort items number 10, 12, and 31. These items, however, were not ranked high in their impact on education.

An examination of Table 2 also reveals that item 7, "a recognized and accepted aspect of teacher education in the future will be that it is non-terminal, continuing, lifelong," was rated the feature or aspect of teacher which would have most impact on education toward the year 2000. Three participants rated it as the one feature or aspect with most impact, 9 participants rated it as the feature or aspect with next most impact, and 11 participants rated it as the feature or aspect with next most impact. The weighting scheme listed provided a weighted total of 3x7, 9x6, 11x5, and so on for a total of 197 (see Table 2).

The ten features or aspects rated as having most impact on teacher education were: lifelong or continuing involvement; closer relationship between theory and practice; inservice teacher education equal in importance to preservice; differentiated programs; more rigorous screening; a greater proportion of total resources to inservice education; candidates prepared to be facilitators of learning; increased emphasis on human relations; candidates spending more time in schools, and increased

emphasis on research in teaching and teacher education. The ten features or aspects with least impact on teacher education were: teacher education in the future will be just about the same as it was in 1975; control will be jointly by universities, professional organizations, employing agencies, and government in that order; control will move out of the hands of the universities; control will be jointly between teacher organizations and teacher education institutions; teacher education will bear little resemblance to teacher education in 1975; the basic programs will vary in length for different candidates; will involve 5 to 8 years of initial preparation; will be much more specialized; will involve more concern and commitment by the total university; and will emphasis the preparation of generalists.

The diversity of opinion among participants is apparent. Only 13 of the 36 features or aspects were not rated by any participant as having “most impact” on education toward the year 2000, and 24 of the 36 were not rated as having “least impact.” Inspection reveals five items (1, 15, 16, 17, and 19) which were rated by participants in every possible way, from most impact to least impact on education. Despite this diversity, the agreement among participants with respect to the features or aspects with most impact is impressive, and perhaps even more so with respect to least impact.

Comparisons

Since this study was designed to replicate, at least in part, a 1970-71 study on the same topic, the first comparison will be with this earlier study.

The basis for choosing participants was identical in each study: the chief administrative officer, and in the event that the institution enrolled more than one thousand students, an additional person for each 1000 students or fraction thereof, in the English-language teacher education institutions in Canada. Table 3 shows the comparisons.

TABLE 3
COMPARISONS OF PARTICIPANTS

	Institutions	Institutions Represented	Individuals Contacted	Participants Who Completed all Parts
1970-71	41	27	55	40
1974-75	39	32	61	41

The overlap of participants after four years was four persons of whom three were chief administrators and one acting for the administrator who was on leave. Turnover is high, holding power is low, drop-outs abound among chief administrators and their designees! From the point of view of the two studies, with a ten percent overlap in participants, there are basically two different sets of participants.

The overlap in statements about the future of teacher education was considerable. In the first study there were 38 statements of features or aspects. While wording varied there were 21 statements common to the two studies. These common statements express the view that there will be a longer

period of preservice preparation; extensive use of internship; many teachers prepared for other than K-12; teacher education just about the same as it is now; teacher education that is lifelong or continuing; more stress on selection with reference to speech, English usage, mental health, and human relations; shared control with teachers' organizations concerning goals, curriculum, recruitment, selection and certification; more specialization; individualization; differentiated programs; more attention to competency based approaches; stress on ethics and values; preparation in the use of educational technology; individualization to the point of no commonality; preparation of teachers as facilitators of learning. It should be noted that in two instances (selection and shared control) the earlier study had several statements expressing different features or aspects while the present study has only one.

Features or aspects contained in the earlier study but not found in the present one follow: Teacher education will be the responsibility of universities; there will be no *institutions* devoted to preparing teachers; teacher education will be involved with constant or periodic re-evaluation of teachers who will have to requalify to retain certification; certification specific to the area of specialization; two major branches of teacher education: educational technology and educational psychology; preparation in working as a leader or member of a group; half "common core" and half specific to specialization re function, level, staff differentiation; general education and subject specialization relegated to the background by psychological, sociological, cultural values, guidance, and similar studies; general methods related to the cognitive disciplines replacing specialized methods; lectures to disappear; professors of education to spend as much time in the field as they do in the university; the professor of education to remain only as a research scholar; disillusionment with innovation and change in teacher education; candidates specialized both in individualization and in group processes; teacher education to be involved in making education an effective force for reducing social and other inequalities.

The features or aspects found in both studies are the ones mentioned most frequently in the initiation of items by participants (Part I), while the differences between the two lists occur in the features or aspects originated by fewer participants.

Some of the features or aspects found in the second study which were not specifically mentioned in the first are: emphasis on humanistic concerns; stronger data base (more scientific); emphasis on research in teacher education; personal development; concern and commitment by the total university; control of teacher education to move out of the hands of the universities; community involvement; a closer relationship between theory and practice; interdisciplinary involvement; preparation and placement consistent with personal characteristics; increased emphasis on controlled clinical experience.

Considering that there was a time lapse and considering that the second study substantially (90%) involved different people, the overlap between the features or aspects mentioned is substantial. Views of the future may change, but in the short run, they retain a considerable stability.

It is not possible to generalize on the basis of the participants common to the two studies, but some interesting comparisons are possible. First,

they tend to be no different than other participants in the study, their initial statements and their reactions to the items in Part II reflecting this. Second, they have generally moved with changes that have occurred since 1970-71 and are aware of current and anticipated thinking and practice with respect to teacher education. Third, they seem more concerned with broad aspects of teacher education than with narrower ones. Fourth, they are consistently in favour of quality programs, which, besides assuming a strong base in knowledge and educational theory and practice, are designed to meet changing times and conditions. Fifth, they recognize the importance of the personal development of the neophyte teacher. In general, they would rate high those items so rated by the participants in this study.

The first study did not ask participants to rate feasibility, desirability, or impact on education, so further comparisons cannot be made on these matters. An interesting task in the first study was that participants were asked to rate when each of the features or aspects of teacher education would be descriptive of teacher education in their own institutions. The finding was that the features or aspects would occur at later dates in their own institutions than in teacher education in general.

A second comparison which can be made is to examine the consistency between forecasts of the future of teacher education, and what people consider to be the goals of education. Lauwerys (1973) stated the conclusions of a 1972 cross-Canada survey of the purposes of education. Included were: diversity is a good thing, the prime function of schools is to convey knowledge and to develop reasoning power, education should aim at helping individuals to enjoy life more richly, e.g., through the arts and by fostering creativity, schools should help children toward getting a job, but, what is even more important, should help them attain self-satisfaction. One can ask whether the future of teacher education should reflect these goals, and whether in our study it did.

It is the view of the authors that forecasts of the future and statements of goals both express participants' *felt needs*. For example, in 1970 not all teacher education in Canada was located in universities or university related institutions. Participants in the study originated this feature or aspect as applying to the future of teacher education toward the year 2000. In 1975, when in fact teacher education in Canada was located in universities, the *felt need* had disappeared and not one of the participants proposed this as a feature of teacher education in the future. If the *felt needs* view is correct, then a one-to-one correspondence between the future of teacher education and the goals of education can be expected only to the extent that the felt needs apply to both.

Perhaps the clearest example is that the persons who responded to the C.E.A. study of Canadian education (Lauwerys, 1973) considered that a prime function of the schools was to convey knowledge and develop reasoning power. No feature or aspect of teacher education in the present study relates directly to this prime goal, presumably because the participants felt that teachers were already, and would continue to be, prepared to promote this goal.

There is consistency, however, between the goals of education and the future of teacher education. For example, each emphasizes diversity, concern for individual development, good human relations, and general

humanistic concerns. These are felt needs; that is, areas for increased emphasis or concern. But there is no one-to-one correspondence between the goals of education (C.E.A.) and the future of teacher education (present study).

Another study, "1,000,000 Opinions: A Report of the Social Values Survey" which Chalmers (1972) conducted in Greater Montreal for the Protestant School Board, found that parents ranked educational goals in this order: vocational preparation, developing self-discipline, helping students understand themselves and others, moral and ethical values, and encouraging originality and creativity. The present study does not mention teacher education to enable teachers to prepare students for the world of work or to get a job, presumably because of the *felt needs* explanation. The other goals, either directly or by implication, are consistent with the preparation of teachers shown in this study.

At the opposite side of Canada, in a study by Coutts, Clarke and Horowitz (1974), in which the goals of education were originated by the panelists, the ten most frequently mentioned were personal development, critical thinking, vocational preparation, universality of educational opportunity, imparting knowledge, constructive use of leisure time, moral and ethical values, lifelong or continuing involvement, ability to cope with change, and developing attitudes favourable to social survival. The study referred to, centered in the province of Alberta, sought the views of a number of groups. Again we find many of the goals of education not specifically covered by the features or aspects of teacher education toward the year 2000 as identified in the present study.

A study by Coutts and Clarke (1972) of the goals of teacher education, also centered in Alberta, which permitted participants to originate the goals, showed that the ten most frequently mentioned were change, values, methods, human relations, inservice education, individualized instruction, selection, facilitation of learning, community involvement, and the learning process. There is much more congruence between these findings and the features or aspects of teacher education toward the year 2000. The notable exception is the strong emphasis on change in the study cited, an emphasis which may have been produced by the procedures used.

It would seem, then, that there is no one-to-one relationship between what various groups see as the goals of education, and our panelists see as the future of teacher education. The authors attribute this in large part to the *felt needs* explanation. This, however, is a matter for some concern and for further study. In the one study cited there was good correspondence between the goals of teacher education as seen by various groups in one province, and the future of teacher education delineated by the participants in the present study.

Discussion

The roles people play and the social setting at a particular time are likely to affect views about the future. This study found that when role was held reasonably constant (the chief administrator in English-language teacher education institutions in Canada) despite the fact that the personnel involved were almost completely different, the view of teacher

education toward the year 2000 remained remarkably constant over a four to five year period. It also identified examples of the social setting at a particular time affecting views about the future, as illustrated in the dropping of the forecast that teacher education would be university based.

The authors have put forward a *felt needs* theory about the generation of statements of goals, or of forecasts about the future in human affairs such as education. The commonplace, the inevitable, the certain, that which is being done and not attracting attention, are the matters which are ignored. The uncommon, the less likely, the uncertain, that which should be done or should be done better, the gap between what *is* and what *ought to be* generate the statements about goals and about the future.

Summary and Conclusions

1. The chief administrative officers in English-language teacher education institutions in Canada when asked to make statements about teacher education toward the year 2000 mentioned most frequently multiple programs, lifelong or continuing involvement, humanistic concerns, an increased emphasis on the school setting, and a stronger data base (more scholarly).

2. The 222 statements originated by participants were classified and 36 representative statements were selected to represent the views expressed. Communication difficulties were illustrated by three of the statements regarded as vague by some participants, e.g., humanistic concerns, facilitators of learning, broader data base.

3. Participants rated the 36 representative statements as to the time when the feature or aspect will be in effect in the majority of teacher education institutions in Canada, its desirability, its feasibility. Many commented on the substance of the statement, explained the reason for their rating, or complained about the communication factor.

4. The features or aspects of teacher education considered to have most impact on education were that in the future it will be lifelong or continuing, will exhibit a closer relationship between theory and practice, will stress in-service teacher education as equal in importance to preservice, will provide differentiated programs, and will involve more rigorous screening.

5. In general, the features or aspects of teacher education toward the year 2000 deemed to have most impact on education were seen as desirable or very desirable, difficult or very difficult. The road is hard!

6. There is substantial agreement with the view of teacher education toward the year 2000 as seen by participants in an earlier (1970-71) study. The participants were substantially different persons holding the same positions.

7. There is no one-to-one correspondence between teacher education toward the year 2000 as seen by the participants in this study and the goals of education as outlined in various studies. This is explained, in part at least, by differences in the *felt needs* of participants at a particular time.

8. There is close agreement between teacher education toward the year 2000 as seen by the participants in this study and the goals of teacher education cited in one previous study.

The study was supported by a grant from the Alberta Advisory Committee on Educational Studies.

References

- Only those references directly pertinent to this study are cited. The general references most often referred to by the authors include Chapter XI, *Teacher Education in the Future*, (1975), Seventy-fourth Yearbook, N.S.S.E., and *The Journal Teacher of Education*, 1974, Vol. 25, No. 2, thematic section on *Futurism in Education*.
- Chalmers, R., & McPhie, G. *1,000,000 opinions: A report of the social values survey*. Montreal: The Protestant School Board of Greater Montreal, 1972.
- Clarke, S. C. T., & Coutts, H. T. The future of teacher education. *The Journal of Teacher Education*, 1971, 22 (4), Winter. Also reprinted in R. W. Hostrop (Ed.), *Foundations of futurology in education*. Homewood, Ill.: ETC Publications, 1973, 114-133.
- Coutts, H. T., & Clarke, S. C. T. *The goals of teacher education*. Research Monograph #20, The Alberta Teachers' Association. Edmonton, Barnett House, 1972.
- Coutts, H. T., Clarke, S. C. T., & Horowitz, M. The goals of education. Unpublished manuscript. The University of Alberta, 1974.
- Lauwerys, J. *The purpose of education: Results of a C.E.A. survey*. Toronto: Canadian Education Association, 1973.

R. O'REILLY

The University of Ottawa

Classroom Climate and Achievement in Secondary School Mathematics Classes

The purpose of this study is to show that the psychological climate of a classroom has an effect on the learning of students. Consequently teachers and supervisors should monitor climate and develop techniques for improving it. The subjects for the study were 1,100 secondary students in 48 mathematics classes in Eastern Ontario. Variables in the study included pupil scholastic aptitude, student achievement, classroom climate as measured by the Learning Environment Inventory (LEI) and pupil characteristics such as social background, liking for school and future orientation. (Dr. O'Reilly is Associate Professor in the Faculty of Education, Graduate Studies, at the University of Ottawa.)

The purpose of this paper is to examine classroom climate as an intervening, or process, variable in education. Administrators and supervisors examine pupil achievement when they assess educational programs. In their evaluation they also allow for the native ability and social background of students. Yet, although they claim to pay a great deal of attention to actual classroom behavior, they fail to examine in a systematic fashion the psychosocial environment of the classroom as a significant intervening variable. It is proposed here that climate is an important variable which should be monitored, manipulated and related to educational outcomes. In order to demonstrate this, data collected in an Ontario study will be utilized to examine the relationship between certain social and personal characteristics of pupils, the climate of grade 9 and 10 mathematics classes, and mathematics achievement of students.

Background

The most persistent questions faced by the educational researcher are those which ask, "What accounts for the variation in the learning of pupils achieved in the classroom?" Thus far, we have been quite successful in relating achievement to student aptitude and to such home factors as socioeconomic status, expectations for success and level of parents' education. Pupil aptitude and scores on I.Q. tests usually account for about 50 per cent in the variance in

school achievement whereas home factors account for 65 percent of the variance in standardized tests. It is on the basis of results such as these that social reformers state that schools do little to reduce effects of social inequalities. Jencks (1972) does suggest in his conclusions, that it is the social environment that immerses the school and the child which accounts for learning rather than curricula, teaching style or the availability of resources. He recommends that schooling be thought of as an end in itself rather than as a means to other ends such as cognitive development or as preparation for adult responsibilities. He goes on to say that although schools should have goals, for all enjoyable activities require purpose, looking back over his data, he concludes that it doesn't matter what those goals are.

Educators must study carefully the reports of studies such as Jencks; however, to accept his conclusions and the model of schooling implied in those conclusions would be to deny our profession as educators. Education is predicated on the premise that, as a result of performing given activities, pupils will change and the change will be in given direction. That is, schools aid children to develop in desired and desirable directions.

We do not deny the existence of heredity. But it is our point of departure; not our end-point.

We agree that the social context of learning is also important; but we deny that the macro-social environment is so overwhelming that the micro-environment of the school and of the classroom is insignificant for learning.

Modern behavioral science, since the time of Kurt Lewin, has been constructed on the dual premise that the social-psychological environment or climate of a work group strongly influences outcomes, and that this climate is influenced by the group leadership and is thus subject to manipulation (Lewin, 1947).

Although a number of methodologies have been devised to research these variables in schools, few studies have been designed to relate climate to learning outcomes. A recent survey located only 13 such studies in addition to a series to be discussed below. The thirteen studies indicated few clear-cut relationships between climate and achievement. The authors attribute this, among other reasons, to the poor validity of instruments, and the use of global concepts of climate rather than breaking it down into its components (Nielsen & Kirk, 1974). In another survey of the research, Rosenshine (1970) encountered similar difficulties.

One classroom climate rating form which has enjoyed some success is the *Learning Environment Inventory* (LEI), developed by Walberg and Anderson (1968) for the evaluation of the Harvard Project Physics. This instrument has been useful in relating climate to achievement; it has also been useful in that teaching methods also related to climate are measured by the LEI.

The final version of the scale contains 105 statements describing typical school classes. Each student expresses his agreement or disagreement on a four point scale. The items are grouped into 15 scales defined as follows:

1. *Cohesiveness*: The feeling of intimacy that has developed as a result of several individuals interacting over a period of time.
2. *Diversity*: The extent to which the class provides for a diversity of pupil interests and activities.

3. *Formality*: The extent to which behavior within the class is guided by formal rules.
4. *Speed*: The rate of progress of the class.
5. *Environment*: The physical environment, including the amount of space available and the type of recreational equipment.
6. *Friction*: The extent to which conflict may affect the behavior of the class.
7. *Goal Direction*: The recognition of goals and their subsequent acceptance by the group.
8. *Favouritism*: The extent to which pupils possess a low academic self-concept.
9. *Cliqueness*: Aims at revealing the extent to which cliqueness exists in a classroom and its influences on social interaction.
10. *Satisfaction*: The extent to which students like or dislike their class.
11. *Disorganization*: The extent to which students consider their class to be disorganized.
12. *Difficulty*: The relative perceived difficulty levels of various courses.
13. *Apathy*: Complements the cohesiveness scale, but also indicates if individuals within the class feel any affinity with class activities.
14. *Democratic*: Indicates the extent to which a “democratic” atmosphere exists within a classroom.
15. *Competitiveness*: The degree of competitiveness existing within the class.

Several reliabilities are available. For class groups, intraclass correlations range from .31 to .92 for each scale. Test re-test correlations for each scale range from .43 to .73. The LEI was successfully utilized in a variety of experimental and correlational studies, described by Anderson and Walberg (1974).

Design of the Study

The unit of analysis in this study is the classroom, and all scores are average classroom scores. Forty-eight classrooms from 12 secondary schools from four boards of education in Eastern Ontario constituted the sample. To fulfill certain requirements of the major project, a combination of random and judgmental sampling was employed. The forty-eight classes were equally divided between the grades 9 and 10 levels (or equivalents) and between advanced academically-oriented and general or terminal classes. A total of 1,100 students participated.

Data Collection

Data were collected from students with the following instruments:

1. *Stanford Achievement Test* (SAT): Mathematics Form W.
2. Biographical data.
 - (a) parents' education
 - (b) dislike for school
 - (c) future planning
3. LEI.
4. *School and College Ability* (SCAT) *Test*, Series II.

Questionnaires and tests were administered in two sittings with each of the 48 classes. One sitting involved the administration of the *Stanford Achieve-*

ment Test—Mathematics—Part A, and the SCAT—Series II, Form 2A in Grade 10 Classes, Form 3A in Grade 9 Classes. The Stanford was administered to 2/3 of the class while the SCAT was administered to the remaining one-third. In the other sitting, the LEI and the Biographical Inventory were administered to one half of the class while other tests were administered to the remainder of the class. Males and females were proportionally represented in each segment of data collection.

The order of administering sets of instruments in the two settings was alternated to eliminate order effects. This method of collecting data was consistent with procedures used by Walberg and Anderson (1968), and was economical in terms of research effort.

Analysis of Data

In order to analyze the results obtained in this project, the framework of INPUT-PROCESS-OUTPUT will be used.

Input variables. Input variables are usually beyond the control of the educator. Those given are: (1) scholastic aptitude (I.Q.) and (2) biographical indices. Correlations with Achievement are shown in Table 1.

TABLE 1
MATRIX OF INPUT AND OUTPUT VARIABLES

Variable	2	3	4	5	R
1. Achievement	.83*	.48*	-.41*	-.26	.88*
2. Scholastic Aptitude	1.00	.45*	-.33*	-.26	
3. Parents' Education		1.00	-.08	-.30*	
4. Dislike for Schooling			1.00	.23	
5. Career Planning				1.00	

* $p < .05$ N = 48

Scholastic aptitude is related to achievement (verbal, $r = .62$; mathematical, $r = .86$). Of the biographical indices, parents' education is related to achievement ($r = .48$); dislike for schooling is similarly related ($r = -.41$). The third biographical variable, labeled "career planning", was intended to obtain an index of the student's future orientation, but it proved to be unsatisfactory in this study.

Taken altogether, these input variables account for 77% of the variance in achievement ($R = .88$).

Process Variables. The process variables are the fifteen scales of the LEI. Taken together, they account for 67 per cent of the variance in achievement ($R = .82$).

The LEI variables are not independent of input variables. This one might expect, as pupils bring a great deal of their world into the classroom. Pupils have a significant effect on the climate of the classroom.

Eleven of the 15 LEI variables are associated with input variables (variables 5, 6, 7, 8, 9, 10, 11, 12, 13, 14 and 15). Scholastic aptitude is associated with

nine LEI scales. All input variables are related to the variables Friction, Satisfaction and Disorganization. In addition, “Dislike for schooling” is related to Favoritism, Apathy and Democratic variables. “Parents’ education” is related to an additional three variables: Goal Direction, Difficulty and Apathy.

TABLE 2
CORRELATIONS OF INPUT-PROCESS AND OUTPUT VARIABLES

LEI Scales	IQ	BI(1)	BI(2)	BI(3)	Achievement
1. Cohesiveness					.30
2. Diversity					
3. Formality					
4. Speed					
5. Environment			-.46		.32
6. Friction	-.54	-.31	.45	.35	-.59
7. Goal Direction		.30			
8. Favoritism	-.42		.45		-.44
9. Cliquesness	-.33				-.34
10. Satisfaction	.43	.31	-.49	-.39	.45
11. Disorganization	-.50	-.35	.48	.48	-.52
12. Difficulty	.40	.31			.33
13. Apathy	-.52	-.40	.42		-.44
14. Democratic	.33		-.41		.47
15. Competitiveness	.35				
R					.82

IQ is SCAT-Total
BI(1) is Parents' Education
BI(2) is Dislike for Schooling
BI(3) is Career Plans.

Only significant correlations are shown. N = 48

Ten LEI variables are related to achievement. Achievement in mathematics classes is characterized by high scores on the following variables: Cohesiveness (.30); Environment (.32); Satisfaction (.45); Difficulty (.33); and Democratic (.47). High achieving classes are noted by low scores on Friction (-.59); Favoritism (-.44); Cliquesness (-.34); Disorganization (-.52); and Apathy (-.44).

Discussion

There are a number of concerns here for the administrator.

If Input is related to both Process and Achievement, can the teacher be in any way held accountable for the classroom climate or for the achievement of pupils?

In order to answer that question, three elements must be considered. First, do input variables account for all the variance? Second, can teachers modify the environment of the classroom? Third, can school be a fruitful experience for all pupils?

One technique available to answer the first question is to compute partial correlations, a technique which first subtracts the common variance which two

variables share with a third variable. Walberg and Anderson (1972) computed correlations between LEI variables and year-end provincial examination scores covering eight subject areas with intelligence partialled out of the relationship. Their sample was 64 grade 10 and 11 classes from eight schools in Montreal. Twelve of the LEI scales were significantly related to overall achievement. The size of the correlations, on the average, were 10% smaller than the correlations between achievement and LEI scales. The data from the O'Reilly and Garland (1974) is less satisfying (Table 3). Correlations between achievement and LEI scales with total SCAT scores partialled out yielded only three significant correlations (Diversity, $-.35$; Friction, $-.30$; Democratic, $.36$). In this case, the Walberg and Anderson data should be given greater credence, since they related their scores to actual examination results, whereas the O'Reilly and Garland study utilized standardized achievement tests which are somewhat confounded with scholastic aptitude tests. The correlation between achievement and aptitude in the O'Reilly and Garland study was $.82$ and in the Walberg and Anderson study, it was only $.38$.

TABLE 3
CORRELATIONS BETWEEN LEARNING ENVIRONMENT SCALES AND
MATHEMATICS ACHIEVEMENT SCORES (SAT) AND PARTIAL COR-
RELATIONS

Environment Scales	Mathematics Achievement	Partial Correlations, Controlling for...			
		Intelligence SCAT-Total	Parents' Education	Dislike for Schooling	Future Plan Orientation
1. Cohesiveness	.30		.29	.25	.30
2. Diversity		-.35			-.26
3. Formality					
4. Speed					
5. Environment	.32		.32		.28
6. Friction	-.59	-.30	-.53	-.49	-.55
7. Goal Direction					
8. Favoritism	-.44		-.38	-.32	-.41
9. Cliqueness	-.34		-.27		-.32
10. Satisfaction	.45		.36	.31	.39
11. Disorganization	-.52		-.43	-.40	-.47
12. Difficulty	.33			.40	.31
13. Apathy	-.44		-.31	-.32	-.41
14. Democratic	.47	.36	.41	.36	.43
15. Competitiveness	.25			.31	
R	.82				

$r_{.05} = .29$ Blanks indicate $r < .25$. $N = 48$

In considering correlations between achievement and aptitude when biographical indices are partialled out, the effects of the LEI scales remain pronounced. When controlling for parents' education, seven of the ten LEI scales remained significantly correlated. Overall the correlations are approximately 10% lower.

When controlling for the pupils' dislike for schooling, eight correlations remain significant.

Controlling for future plan orientation produces almost no changes.

Thus, although students' characteristics do predict class environment to some extent, there is still room for the effect of teacher and school controlled variables upon academic achievement.

The answer to the second question cannot be answered by the data presented here, but there is the inference that teachers can influence climate and consequently achievement. Further, it is axiomatic in the literature on leadership, supervision and management, that leaders do have an impact on both group climate factors and on achievement (Likert, 1967). Early studies by Anderson and Brewer (1945) established that a teacher can have an effect on classroom climate very early in the school year. It would seem logical to assume that the climate of a classroom is the result of the interaction of pupil and teacher characteristics and that the influence of the teacher would not be insignificant. Consequently a study of classroom climate early in the school year would be advisable.

To the third question, "Can the school be a fruitful experience for all students?" the reply must hinge on the level of ability of the teacher. In an age when teachers were less well trained and when there were few opportunities for professional development, it could have been expected that teachers would easily teach the able and motivated student, whereas the others, subjected to repeated failure, would eventually drop out. Today it is difficult to excuse a university trained profession with ample opportunities for professional development, for failure to provide the majority of all pupils with satisfying learning experiences within the school.

The identification of important aspects of climate gives an added tool to teachers and supervisors. In management theory, it has been shown that trends in the change of direction in climate can be detected fairly early, whereas the effects of such changes on productivity might be delayed by as much as two years (Likert, 1967). Educators could begin to systematically monitor classroom climate and attempt to influence its nature. Expertise to do this is rapidly appearing in the education profession (Phi Delta Kappa, 1974). A supervisory program could be designed which would call for early monitoring of classroom climates, and if the situation warrants it, expert assistance, perhaps in the form of a school guidance teacher or an organizational development professional, would assist in a program of climate change.

To return to the discussion of the study by Jencks, on the basis of the material presented in this paper, we can conclude:

1. Classroom climate is a significant factor in pupil achievement.
2. Although pupil personal and social characteristics are important correlates of achievement, climate in some instances is a more important factor.
3. Although input, process and output are all intercorrelated, process variables have an independent effect on achievement.
4. Education includes psycho-social factors as well as academic factors and the two are related.

References

- Anderson, G., & Walberg, H. J. Learning environments. In H. J. Walberg (Ed.), *Evaluating educational performance*. Berkeley: McCutchan Publishing, 1974, 81-98.
- Anderson, H. H., & Brewer, J. E. Studies of teacher classroom personalities. *Applied Psychology Monograph*, 1945, 6.
- Jencks, C., et al. *Inequality: A reassessment of the effect of family and schooling in America*. New York: Basic Books, 1972.
- Lewin, K. Frontiers in group dynamics. *Human Relations*, 1947, 1, 5-41.
- Likert, R. *The human organization*. Toronto: McGraw-Hill, 1967.
- Nielsen, H. D., & Kirk, D. H. Classroom climates. In H. J. Walberg (Ed.), *Evaluating educational performance*. Berkeley: McCutchan Publishing, 1974, 57-79.
- O'Reilly, R., & Garland, P. *Paradigm for evaluation in the high school*. Ottawa: University of Ottawa, 1974.
- Phi Delta Kappa. *School climate improvement*. Bloomington, Ind.: Phi Delta Kappa, 1974.
- Rosenshine, B. Evaluation of instruction. *Review of Educational Research*, 1970, 40, 279-300.
- Walberg, H. J., & Anderson, G. Classroom climate and individual learning. *Journal of Educational Psychology*, 1968, 59, 414-419.
- Walberg, H. J., & Anderson, G. Properties of the achieving urban classes. *Journal of Educational Psychology*, 1972, 61, 381-385.

D. F. ROBITAILLE

University of British Columbia

Classroom Personality Patterns of Teachers of Secondary Mathematics

The 167 teachers of secondary mathematics in a public school system in Montreal, Quebec were rated by four judges working independently of one another with usable data being obtained on 127 of the teachers. Each teacher was assigned the sum of his four ratings as an effectiveness score. The thirty teachers with the lowest effectiveness scores were designated as minimally effective and the twenty-five with the highest scores as highly effective.

The secondary school principals judged the teachers' classroom personality patterns by assigning each teacher to one of six personality categories. Three of these categories had positive denotations and three had negative denotations. Results showed that all twenty-five effective teachers were assigned positive categories with over 80% being rated as responsible, businesslike, and systematic in their teaching. Over 60% of the ineffective teachers were assigned to negative categories with most of them being rated as dull and routine teachers. (Dr. Robitaille is Assistant Professor in the Faculty of Education, The University of British Columbia.)

The area of teacher effectiveness has produced probably more research than any other in the field of education. In spite of all this effort, until quite recently writers in the area have been of the opinion that very little was known for certain about the ingredients constituting effective teaching or about the relationship between teacher personality and teaching effectiveness (Getzels & Jackson, 1963).

It may be the case that one of the reasons for the lack of compelling findings from teacher effectiveness research has been that the research questions have frequently been too broad. Much previous research seems to be based on the assumption "... that there is some magic variable that applies to all of teaching, for all students, at all grade levels, in all subject matters, and for all objectives" (Gage, 1972).

It is probably not possible to define and characterize teaching effectiveness in general, at least until a series of preliminary questions has

been answered. Some such preliminary questions might address specific aspects of the teaching act. Gage, for example, has discussed the use of "micro-criteria".

Rather than seek criteria for the overall effectiveness of teachers in the many, varied facets of their roles, we may have better success with criteria of effectiveness in small, specifically defined aspects of the role. (Gage, 1972, p. 95)

These aspects of the teaching act might include such "technical skills" as explaining, questioning, obtaining student participation, providing feedback, and so on.

A second set of preliminary research questions, not necessarily completely disjoint from the first, might address the area of grade level and subject matter differences in teaching. Thus, it may be fruitful to attempt to characterize the effective teaching of secondary mathematics, for example, before attempting to describe effective teaching in general. In other words, it may be the case that the characteristics of effective teachers of a given discipline at a given grade level are quite different in many ways from those of effective teachers of the same subject at a different grade level or from those of effective teachers of a different discipline at any grade level.

There is some evidence to show that teachers of a given subject at a given grade level do exhibit personality characteristics which differ from those of other teachers at that grade level in other subject areas. In the Teacher Characteristics Study (Ryans, 1960), three major patterns of teacher classroom behavior were identified. These were:

TCS Pattern X: warm, understanding, friendly versus aloof, egocentric, restricted teacher classroom behavior;

TCS Pattern Y: responsible, businesslike, systematic versus evading, unplanned, slipshod teacher classroom behavior;

TCS Pattern Z: stimulating, imaginative versus dull, routine teacher classroom behavior.

Ryans' results indicate that "superior" teachers of secondary mathematics and science score significantly higher than "poor" teachers of these same disciplines on all three of the TCS patterns: X, Y, and Z. The teachers who participated in this aspect of the Teacher Characteristics Study were nominated by their respective principals as being outstandingly superior or poor (Ryans, 1964).

The present paper is a report of an attempt to discover which one of the TCS patterns best describes the classroom behavior of a highly effective teacher of secondary mathematics, and which one best describes the classroom behavior of a minimally effective teacher of secondary mathematics. Data such as these may be useful in determining some of the parameters governing teacher effectiveness in the area of secondary school mathematics.

Method

All one hundred sixty-seven teachers of secondary school mathematics, grades 8-12, in a public school system in Montreal, Quebec were rated by each of four judges working independently of each other. The raters were the Assistant Director of Secondary Education, the two Mathematics Coordinators for the system (including the author), and the principal of the

school where the teacher being evaluated taught. Thus each teacher was evaluated by his or her own principal as well as by three persons who had system-wide responsibilities.

The evaluators rated each teacher on a five-point scale ranging from 1 to 5. A score of 1 indicated minimal teaching effectiveness, while a score of 5 indicated a very high degree of teaching effectiveness. No criteria of teaching effectiveness were provided for the raters. They were merely asked to provide their personal opinions of each teacher's effectiveness by assigning a score of between 1 and 5. When the ratings had been completed, each teacher was assigned the sum of his four ratings as an effectiveness score. The lowest possible effectiveness score was 4 and the maximum was 20. Of the total of 167 teachers, usable data were obtained on 127; that is, approximately 76% of the total. The remaining 40 teachers were new to the system and, in each case, at least two of the raters felt unqualified to rate them.

In a related study (Robitaille, 1975), those teachers with effectiveness scores in the range 16-20 were designated as highly effective teachers and those whose effectiveness scores fell in the range 4-9 were designated as minimally effective. Analysis of background data and the results of paper and pencil tests administered to these teachers strongly supported the evaluations given by the raters. That is, the effective teachers were significantly better than the minimally effective teachers on several measures. The testing instruments used for this purpose included the Contemporary Mathematics Test (Massie, 1967), the Teaching Situation Reaction Test (Duncan & Hough, 1966), the Minnesota Teacher Attitude Inventory, and the Purdue Teacher Opinionnaire. Similarly, data obtained from classroom observations of these same teachers produced significant differences in favor of the teachers rated as highly effective with respect to the extent of student participation in their classes.

At the same time as the principals rated each of their teachers, they were asked to select from a given list the personality pattern which best described that teacher's classroom behavior. For this purpose, each of the three TCS patterns was divided into two parts, one with a positive denotation and the other with a negative one. The six resulting categories were: XA—warm, friendly, understanding teacher behavior; XB—aloof, egocentric, restricted teacher behavior; YA—responsible, businesslike, systematic teacher behavior; YB—evading, unplanned, slipshod teacher behavior; ZA—stimulating, imaginative teacher behavior; ZB—dull, routine teacher behavior.

Results and Discussion

The distribution of effectiveness scores is shown in Table 1. Thirty teachers out of the total pool of 127 were rated as minimally effective; that is, their effectiveness scores fell within the range 4-9. Twenty-five teachers were judged to be highly effective; that is, their effectiveness scores fell within the range 16-20. The mean effectiveness score was 12.5 and the standard deviation of the scores was 3.6.

A chi-square test for goodness of fit supported the hypothesis that the distribution of effectiveness scores was not significantly different from a

TABLE 1
DISTRIBUTION OF EFFECTIVENESS SCORES

Score	Frequency	Score	Frequency
4	1	13	14
5	3	14	9
6	2	15	17
7	6	16	6
8	7	17	10
9	11	18	2
10	4	19	5
11	12	20	2
12	16		

N = 127

normal distribution. In applying this test, the three lowest categories (4, 5, and 6) were grouped together because of the small number of cases involved. Categories 18, 19, and 20 were also grouped together for the same reason.

A coefficient of inter-observer reliability was computed using a technique suggested by Kerlinger (1964). This involves the use of a two-way analysis of variance (see Table 2) in order to compute the required value.

$$r = 1 - \left(\frac{V_e}{V_{ind}}\right) = 1 - (0.43/6.56)$$

V_e indicated "the estimate of the error variance" and V_{ind} denotes "the variance resulting from individual differences." This technique produced a 0.94 coefficient of inter-observer reliability indicating a high degree of agreement among the four evaluators.

TABLE 2
ANALYSIS OF VARIANCE OF EFFECTIVENESS SCORES

Source of Variation	Degrees of Freedom	Sums of Squares	Mean Squares
Teachers	126	407.92	3.24
Effectiveness Scores	3	19.69	6.56
Residual	378	161.81	0.43

The findings regarding the teacher's classroom personality patterns as judged by their principals are summarized in Table 3.

The twenty-five teachers who formed the pool of highly effective teachers were all rated by their principals as belonging to one of the three positive categories XA, YA, and ZA. Of the 25, 80% were judged to exhibit responsible, businesslike, and systematic behaviour (category YA) as the predominant trait in their teaching.

Of the thirty teachers judged to be minimally effective over 60%, 19 out of 30, were seen by their principals as exhibiting one of the three negative

TABLE 3
MATRIX OF EFFECTIVENESS SCORES VERSUS PERSONALITY PATTERNS

Personality Patterns	Low (4-9)	Moderate (10-15)	High (16-20)	
XA	10	25	6	41
XB	4	3	0	7
YA	1	30	16	47
YB	4	0	0	4
ZA	0	5	3	8
ZB	11	9	0	20
	30	72	25	

personality patterns in their teaching. More than half of these, 11 out of 19, were seen as being dull and routine in their teaching. Of the eleven minimally effective teachers who were awarded a positive personality pattern over 90%, 10 out of 11, were judged to be warm, friendly and understanding in their teaching. It is possible that some of these teachers were judged to have an excess of personality pattern XA and for this reason were given low effectiveness scores. Also noteworthy is the fact that only 6 of the 25 most effective teachers were placed in category XA.

Two categories, namely YA (responsible, businesslike, and systematic) and ZB (dull and routine), together include more than 50% of the teachers, 67 out of 127 to be precise. It may be that the distinction between these two categories is not as clear as it could be and that some principals had difficulty differentiating between the two. In other words, an evaluator might find it difficult to differentiate between a teacher whose teaching is "routine" and one whose teaching is "systematic" or "businesslike". If it is indeed the case that there is an ambiguity here and that the distinction between the two categories is vague, then these principals may be saying that the majority of these teachers of secondary mathematics exhibit classroom behavior which could be characterized as dull and routine.

Of the three positive categories, ZA (stimulating and imaginative teacher behavior) was utilized least. Only eight teachers, approximately 6% of the total, were seen as being predominantly stimulating and imaginative in their teaching behavior. More interestingly, of these eight only three were rated among the highly effective teachers. This finding seems to indicate that, in the opinion of principals, very few teachers of mathematics are predominantly stimulating and imaginative in their teaching and also that stimulating and imaginative teachers are frequently *not* among those judged to be the most effective teachers.

From another point of view, the finding is that only 3 of the 25 most effective teachers are judged to be stimulating and imaginative. This does not necessarily mean that the other 22 highly effective teachers are not stimulating and imaginative, but rather that one of the other two personality patterns XA and YA is predominant in their teaching behavior.

Still, the result does seem to say that principals find very few of their teachers of mathematics to be particularly stimulating and imaginative in their teaching. Such a finding should be of interest to those who are involved in the training of prospective teachers of secondary mathematics as well as to those involved in in-service programs for teachers.

Summary

The overwhelming majority of the most effective teachers were judged by their principals to be responsible, businesslike, and systematic in their teaching of mathematics. The other two positive categories, XA and ZA, included only 20% of the highly effective teachers.

The results for the minimally effective teachers are not as clear cut. One third of these teachers were rated as warm, friendly and understanding, a positive category, while another 11 of the group of 30 were seen as exhibiting dull and routine teaching behavior, category ZB.

References

- Biddle, B. J., & Ellena, W. J. (Eds.) *Contemporary research on teacher effectiveness*. New York: Holt, Rinehart and Winston, 1964.
- Duncan, J. K., & Hough, J. B. Technical review of the teaching situation reaction test. Unpublished manuscript, The Ohio State University, 1966.
- Gage, N. L. *Teacher effectiveness and teacher education: The search for a scientific basis*. Palo Alto, Calif.: Pacific Books, 1972.
- Getzels, J. W., & Jackson, P. W. The teacher's personality and characteristics. In N. L. Gage (Ed.), *Handbook of research on teaching*. Chicago: Rand McNally, 1963.
- Kerlinger, F. N. *Foundations of behavioral research*. New York: Holt, Rinehart and Winston, 1964.
- Massie, R. O. The construction and use of a test to evaluate teacher preparation in modern mathematics. Unpublished doctoral dissertation, University of Nebraska, 1967.
- Robitaille, D. F. Criteria for assessing the effectiveness of teachers of secondary mathematics. *Journal for Research in Mathematics Education*. 1975, 6, 77-87.
- Ryans, D. G. *Characteristics of teachers*. Washington: American Council on Education, 1960.
- Ryans, D. G. Research on teacher behavior in the context of the teacher characteristics study. In B. J. Biddle & W. J. Ellena (Eds.), *Contemporary research on teacher effectiveness*. New York: Holt, Rinehart and Winston, 1964.

J. P. KLEIN

University of Toronto

Socratic Dialogue vs. Behavioural Practice in the Development of Coping Skills

An attempt was made to increase the competency skills of 44 male students in grades 7 to 9 through the Socratic teaching of the theory of security and independence, the behavioural practice of coping skills, or both. At the end of six half-hour sessions no significant improvement was found in math grades, but directional and significant differences demonstrated that the groups receiving dialectic training increased confidence in their problem solving capacities, and showed a heightened sense of security and self-directedness. (Dr. Klein is with the Institute of Child Study, Faculty of Education at the University of Toronto.)

With the increasingly widespread recognition that the results of traditional forms of psychotherapy with children are a serious disappointment to our optimistic imagination (Levitt, 1963), some clinicians emphasize the need for experimental modes of treatment (Lazarus, 1971) while others shift the entire focus from remediation to prevention (Kohlberg, 1972). Accompanying the latter thrust of inquiry is the substitution of cognition for affect as the major variable of interest. Disruptive emotions are viewed as products of irrational thinking (Ellis, 1962) or maladaptive coping responses (Lazarus, 1966). The road to health is the one which secures ego strength; the child possessing those adaptive skills which lead to the mastery of life stress situations will be free from psychological problems. He has ego strength. Those possessing these skills in smaller degrees are increasingly inclined to utilize inadequate acts of mastery called defenses; they employ no strategies which lead them on to depression; or they act with strategies curiously inappropriate to the stress event called psychosis.

While many investigators explore the life-history antecedents and personality correlates of coping skills, to the author's knowledge this is the first study that attempts to test the hypothesis that ego strength can be experimentally developed.

The first method chosen is the Socratic dialogue. While the juxtaposition of seemingly contradictory arguments in order to stimulate debate leading towards a truthful conclusion has a long classroom history, the method and its purpose as stated by Socrates (Plato, 1927) have never been accurately followed. Socrates never taught before he exhorted and provoked his student into a realization of his ignorance about something of vital importance to his life. Modern psychology has elaborated on the extent to which ego-involvement raises the attractiveness of a task (Ausubel & Sullivan, 1970). In addition, Socrates never used the techniques as a means of communicating factual information. Education was a way of bringing out qualities of character already possessed but that lay dormant and undifferentiated within a person. Socratic dialogue was a way of becoming somebody, not knowing something.

The second method is a behavioural approach which presents a good contrast to much of the theoretical thinking about ego strength. By getting a student to perform directly the final mastery behaviours, we shift from an analysis of the intervening variables that modulate coping skills as they are explained by the ego psychologists (White, 1974), to a simple habitual exercise of those terminal acts which constitute adaptation to life problems.

Method

Subjects

All male students in grades 7 to 9 of a Toronto junior high school were told that for their voluntary co-operation in an experimental project, they would receive their choice of a gift worth five to ten dollars. The mean age of the 52 volunteers was 13.0 years. The students were ordered alphabetically by last name. Blocks of four students were then taken, and a table of random numbers was used to assign the students within blocks to the experimental treatments. Each of the four groups had 13 Ss. The author served as *E*.

Design and Procedure

The four treatment conditions were (a) Socratic learning of the Theory of Security, (b) behavioural training in coping skills, (c) equal amounts of training in (a) and (b), (d) a no treatment control group. Each treatment group met every third school day for 30 minutes during school hours over the span of a month. This resulted in a total of six half-hour sessions for each treatment group. Two one-hour testing sessions were held for all Ss: one directly prior and one directly following the treatment program.

Socratic Dialogue

The Institute of Child Study Security Test (Grapko, 1957) called "The Story of Jimmy" is a written projective measure in which the *S* ranks five possible solutions to a given problem in the life of Jimmy. The alternatives reflect varying degrees of maturity in the manner of accepting consequences for one's own decisions. Those six problem situations most closely related to school life were chosen (numbers 1, 4, 5, 8, 9, 10). For each session the Ss read one of the problems to themselves and ranked the given choices from the one most to the one least likely for Jimmy to choose. The

E, following the Socratic method as outlined in Oliver and Shaver (1966), then called upon an *S* to state his lowest ranked choice and give the reason for his selection. *E*, assuming the Socratic adversarial role, questioned the *S* by presenting a parallel but slightly different situation to which the *S* would have responded differently. The *S* was then asked to state the principle which would lead him to a different decision in each of the two parallel cases. Other *Ss* often intervened to present their different choice for lowest rank, or to criticize the principle enunciated by the first *S*. *E*, interjecting statements about the seriousness of being correct in matters of personal responsibility and the importance of acting like an adult, always concluded by teaching the principle which security theory attributed to that rank. Then the group went on to the second from last choice and proceeded in the above pattern until all five choices were discussed in the half hour session. The six situations thus occupied the six treatment sessions.

Behavioural Training

The *E* requested the *Ss* to list those aspects of academic life for which they were generally held responsible. *E* added those omitted by *Ss* so that the list included controlling academically irrelevant behaviour, paying attention in class, raising your hand to ask questions when unclear about the lesson being taught, doing the homework regularly and independently, and studying before an exam. *Ss* volunteered responsibilities often evaded at home such as straightening the bedroom, dumping the garbage or the brief caring for a pre-school sibling. To insure confidentiality and honesty at the beginning of a session, each *S* would write on a sheet of paper his name and those responsibilities which he had performed since the last session. The *E* requested *Ss* to report verbally any difficulties they had in fulfilling these responsibilities and suggested ways of overcoming them. *E* suggested the four times per week math class as a focus for examining barriers to achieving the behavioural goals.

Combined Condition

Ss in this group spent the first half of each session at the Story of Jimmy and the second half at behavioural training. They covered the same ground as the preceding groups, but learned it less intensively.

Dependent Variables

With the exception of math grade, the following measures were administered to all *Ss* directly before and once again directly after the treatment program.

Anxiety Scale. The Children's Manifest Anxiety Scale (CMAS) (Casteneda, McCandless & Palermo, 1956) was used as an overall measure of anxiety and in security.

Problem Solving. A Child Attitude Inventory for Problem Solving (CAPS) developed by R. S. Crutchfield and M. L. Covington (Johnson & Bommarito, 1971) is a 33 item questionnaire measuring beliefs about problem solving in general and one's own ability at problem solving in particular.

Responsibility for Achievement. The Intellectual Achievement Responsibility Questionnaire (IAR) (Crandall, Katkovsky, & Crandall, 1965) is a

34 item questionnaire measuring the extent to which a child attributes his intellectual progress to himself or to external factors beyond his control. Seven scores can be derived by dividing the items into those mentioning behavioural effort (e) and those not mentioning a specific cause (u for undifferentiated), and events with positive (+) or negative (-) outcome. Thus I+e is arrived at by adding the S's internal responses (I) to items in which the positive events (+) stem from an internal response indicating behavioural effort (e). Similarly, scores for I+u, and I-u are calculated. The two I+ subsets can be added to give an additional score as can the two I-subsets; and an overall score can be obtained by adding these two. Every scored item raises the S's measure of inner-directedness.

Math Grade. By prior arrangement with the math teacher but unknown to the Ss, a new unit of math study began with the initiation and ended at the conclusion of the treatment program. Marks in math on the Fall report card were the pretreatment measure of academic achievement, and an objective quiz on the unit, administered and scored by the teacher, served as the post-treatment measure.

Results

Three Ss in Group 1, four Ss in Group 3, and one S in Group 4 were excluded from the study because they either completed the tests incorrectly or missed more than two treatment sessions. For the remaining 44 Ss, an analysis of covariance was performed using the pretreatment measure as covariate and post-treatment measure as criterion for each of the dependent variables. All tests were nonsignificant with the following exceptions: a directional effect was evident on the CMAS, on the CAPS, and on the Iu+ (see Table 1). The adjusted criterion group means are presented in Table 2.

TABLE 1
ANALYSIS OF COVARIANCE

Source	Sums of Squares	df/df	F
Children's Manifest Anxiety	139.42	3/39	2.29*
Attitude for Problem Solving	815.07	3/39	2.31*
IAR (Iu+)	12.74	3/39	2.56*

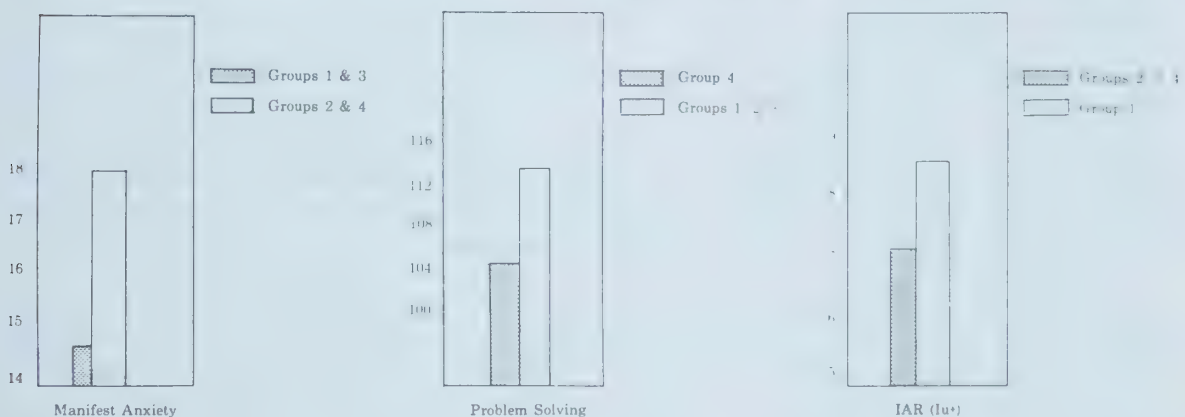
* $p < .10$

TABLE 2
ADJUSTED CRITERION GROUP MEANS

	Group 1	Group 2	Group 3	Group 4
Children's Manifest Anxiety	14.65	18.58	14.44	17.66
Attitude for Problem Solving	115.96	110.95	112.40	104.02
IAR (Iu+)	8.49	7.07	6.98	7.35

A Duncan's Multiple Range test was performed for each of these three variables to identify the source of differences between groups. On the CMAS, Groups 1 and 3 formed a homogenous subset as did Groups 2 and 4. The former subset showed reduced anxiety. The subsets differed directionally ($p < .10$) at their closest points (Groups 1 & 4), and significantly ($p < .05$) at their most distant points (Groups 2 & 3). An analysis of the CAPS revealed one homogenous subset: Groups 1, 2, & 3 not differing amongst themselves, but Group 4, scoring the lowest for problem solving attitude differing significantly ($p < .05$) from Group 1 and directionally ($p < .10$) from Group 2 and from Group 3. An analysis of the IAR (Iu+) showed one homogenous subset: Groups 2, 3, and 4 not differing amongst themselves, but Group 1, showing the highest rating for academic responsibility, differing significantly ($p < .05$) from Group 2, 3, and 4. These effects are presented in Figure 1. Group means of homogenous subsets are averaged for easy identification.

FIGURE 1
A COMPARISON OF GROUP MEANS FOR THREE DEPENDENT VARIABLES



A Pearson Product-Moment correlation was computed between the pretests of all the dependent variables and the pretest to post test difference scores for these variables. Only those correlations relevant to the significant dependent variables and having a probability level of at least .05 are reported. In general these moderate correlations ranged from .30 to .50. Anxiety was negatively correlated with IAR positive (+) scores and with CAPS scores. Iu+ was positively correlated with other IAR scores and with CAPS scores. With reference to change scores, lesser anxiety accompanied increased scores on the CAPS. Increments in Iu+ were accompanied by increments in CAPS scores. No math measures correlated significantly with any other score.

Of the written reports for the performance of behavioural tasks of responsibility handed in by Groups 2 + 3, over 90% claimed successful completion of the tasks. But the language employed by the Ss so closely resembled the phrases used during the treatment session that their validity is suspect. The Ss' verbal reports about problems encountered in the completion of given tasks indicated a genuine involvement with the desired goals.

Discussion

Emergent from the pattern of results is the conclusion that the Socratic dialectic study of the "Story of Jimmy" is the most potent element in building coping skills. Its intensive study in Group I provides the first evidence that locus of reinforcement can be experimentally altered towards greater self-reliance. Its efficiency in strengthening security (reducing anxiety) appears in groups receiving smaller (Group 3) and greater (Group 1) amounts of study. Along with behavioural practice, such study also increases the self estimate of problem solving capacity and decreases the subjective estimate of the difficulty in solving problems in general.

If adaptation is accepted as the global concept under which various coping and mastery skills can be subsumed, a taxonomy of ego strength may be expounded. Following Piaget's developmental notion of adaptation as a form of compromise between the individual and his environment, a minimal condition necessary for successful adaption is the accurate understanding of the nature of the person-environment conflict which has produced the stress. It is hard to imagine a child making an adept adjustment to conflict if the parameters are unspecified or unclear to him. Secondly, the resolution of the stress event often assumes that the individual must at times relinquish self-direction in favour of being controlled by some external force. At this point the assimilation-accommodation equilibrium is upset by the predominance of the latter. Optimally speaking, adaptation must include a compromise in which the individual maintains a balance between freedom of action and compliance as expressed in the ability to use help.

In this light it is easy to see why the "Story of Jimmy" is so excellent a device for building (as well as measuring) adaptation. The various vignettes specify a conflict situation between the child and his environment. Each of the alternate solutions offered reflect one of the five levels of independence that graduate from independent security (assimilation of the environment to current cognitive schema) through mature dependent security (sharing the consequences of decisions) to insecurity (surrender of autonomy.) Far from emitting a negative reaction to being pushed beyond their present level of security, the students relished the understanding of social requirements for mature adult behavior. Maybe Socrates was right in assuming the existence of ego strength which only needed cultivation. This cognitive growth justifies the admission of intervening variables as a necessary component of any short-term program for enhancing adaptability to stress. Behavioural practice may develop such intellectual strength (all treatment groups showed gains on CAPS) but only in a retrospective trial and error attempt to formulate guiding principles. Security theory presents them at the outset.

How can we conclude without paying our respects to the situational complexities of human behaviour? Lazarus (1974) claims that one needs to gauge the entire situation before evaluating the level of coping efficiency. A common response to stress is an increase in affiliative behaviour. The students in the study recognized this difficulty when they claimed that less mature forms of security such as begging to be forgiven were generally more effective with teachers. Schools preach and practise the ethic that the meek shall inherit the earth, a maladaptive gesture in the business world.

Situational variance may account for the negative results with math grades. Perhaps previous knowledge of math is a stronger determinant of progress than any coping device. Future work should attempt to devise sound behavioural measures of adaptability. Lazarus (1974) has already warned us away from global measures, having discovered that the parts of an overall coping behaviour may serve antithetical functions. Only longitudinal data will ascertain whether gains in adaptability are permanent, and whether Kohlberg (1972) is correct in hypothesizing that the presence of various forms of competence is the best predictor for the absence of mental illness in adults.

This study was supported by the Ontario Educational Research Council.

References

- Ausubel, D. P., & Sullivan, E. V. *Theory and problems of child development*. New York: Grune & Stratton, 1970.
- Castaneda, A., McCandless, B. R., & Palermo, D. C. The children's form of the manifest anxiety scale. *Child Development*, 1956, 27, 317-326.
- Crandall, V. C., Katkovsky, W., & Crandall, V. J. Children's beliefs in their own control of reinforcement in intellectual-academic achievement situations. *Child Development*, 1965, 36, 91-109.
- Ellis, A. *Reason and emotion in psychotherapy*. New York: Lyle Stuart, 1962.
- Grapko, M. F. The Institute of Child Study Security Test, Elementary Form—*The story of Jimmy; Manual*. Toronto: University of Toronto Press, 1957.
- Johnson, O. G., & Bommarito, J. W. *Tests and measurements in child development: A handbook*. San Francisco: Jessey-Bass, 1971.
- Kohlberg, L. The predictability of adult mental health from childhood behavior. In B. B. Wolman (Ed.), *Manual of child psychopathology*. New York: McGraw-Hill, 1972.
- Lazarus, A. A. *Behaviour therapy and beyond*. New York: McGraw-Hill, 1971.
- Lazarus, R. S. *Psychological stress and the coping process*. New York: McGraw-Hill, 1966.
- Lazarus, R. S., Averill, J. R., & Opton, E. M., Jr. The psychology of coping: Issues of research and assessment. In G. V. Coelho (Ed.), *Coping and adaptation*. New York: Basic Books, 1974.
- Levitt, E. E. Psychotherapy with children: A further evaluation. *Behaviour Research & Therapy*, 1963, 1, 45-51.
- Oliver, D. W., & Shaver, J. P. *Teaching public issues in the high school*. Boston: Houghton Mifflin, Co., 1966.
- Plato, *Selections*. R. Demos (Ed.) New York: Scribner's, 1927.
- White, R. Strategies of adaptation: An attempt at systematic description. In G. V. Coelho (Ed.), *Coping and adaptation*. New York: Basic Books, 1974.

E. D. HOBBS

University of British Columbia

Methodological Problems in Conservation Testing With Particular Reference to Volume Conservation

Methodological problems associated with conservation testing, related to task validity, horizontal decalage, and inference criteria, are isolated and discussed. Conservation tasks generally are analyzed, and four elements essential to task validity are reduced. On the basis of the theoretical discussion, an unequivocal testing procedure for volume conservation employing nonverbal responses and objective criteria is designed. (Dr. Hobbs is Assistant Professor in the Department of Science Education, University of British Columbia.)

The ultimate concern in this article is the design of a valid, objective test of volume conservation. First, a number of methodological problems endemic in Piaget-related research will be isolated and discussed. Conservation tasks specifically will then be considered and finally, a reliable procedure for testing volume conservation, using valid objective criteria, will be proposed.

The central developmental predictions of Piaget's theory concern sequences and synchronisms. That is, certain abilities are supposed to emerge in sequences, while others are supposed to emerge concurrently. The theory posits that these sequences and synchronisms are not coincidental, do not imply innate structures, and do not result purely from learning experiences. Instead, they are viewed as the inevitable result of certain laws of development, with development seen to be the progressive construction of equilibrated systems of operations. The nature of the theory being so, related research usually involves one or more of four kinds of activity: (a) classifying subjects according to developmental stage or "presense" of a particular ability. (b) empirically verifying the sequences and synchronisms, (c) setting up scales of cognitive growth on the basis of the sequences, and (d) articulating the stage theory.

The first kind of activity, classifying subjects according to developmen-

tal level or performance on Piagetian tasks, may be done in order to characterize some defined population, or to compare populations in cross-cultural studies.

The second kind of activity, verifying invariant sequences or synchronisms, is necessary to establish the empirical base of Piagetian theory. According to Brainerd (1973), "it is the hypothesized *correspondences* and *sequences* which are the crucial empirical predictions [of Piagetian theory]." Studies involved in this kind of activity usually require the classifying of subjects in cross-sectional or longitudinal surveys. Some employ training experiments, which also involve classifying subjects according to performance on Piagetian tasks, after and often before training procedures.

The third kind of activity, setting up scales of cognitive growth, as exemplified by the work at Temple University (McLaughlin, Bentler, & Stephens, 1970), involves assuming that invariant developmental sequences exist. A large number of Piagetian tasks are standardised and placed on an ordinal scale. A subject's position on the scale may then be determined by the classification of his behavior on the tasks.

The fourth kind of activity, articulating the stage theory, is necessary because so far the sequence of development is worked out only in coarse detail. The fine structure remains to be established (e.g., Somerville, 1974). Another problem of articulation, one with which the Genevan group is occupied, is that of the mechanism by which subjects pass from one developmental stage to the next (Piaget, 1971). Both kinds of problem may require training experiments, which in turn involve classifying subjects.

Thus all four streams of theory-related research involve classifying subjects according to performance on Piagetian tasks. All Piagetian tasks may be reduced to the following simple paradigm: (a) *S* is exposed to a complex set of stimuli—"the task," (b) *S*'s behavior is observed, and (c) from this behavior, developmental stage, psychological processes, or thought structures are inferred in *S*. Three fundamental methodological problems arise in this paradigm. The first may be termed "task validity," the second is concerned with the phenomenon of horizontal decalage, the third is that of inference criteria. These will now be discussed *seriatim*.

Task Validity

The term "validity," applied to a Piagetian task, has two meanings. The first refers to the validity of the Piagetian interpretation of subjects' performance on the task. The second refers to the validity of applying that same interpretation to a variation of the task. The act of labelling a particular task as, for example, a volume conservation task, is actually an interpretation of subjects' behavior on the task. To give the same label to a variation of the task is to make a judgment about the equivalence of the task and variation. Such a judgment may not be valid because subjects' behavior on the variation may have a different interpretation in the same theory. Thus the designing of a task necessarily involves interpretation of the theory. If the experimenter is assuming the validity of the Piagetian interpretation of a given task, he must be careful not to apply that interpretation to a variation of the task without justification in terms of the theory. Classification of subjects in experiments which lie within the framework of the theory

will have little meaning if based on tasks which in effect reinterpret the theory. On the other hand, if the experimenter is challenging the validity of the Piagetian interpretation of a task, he must be careful not to apply the Piagetian interpretation to a variation of the task, unless that interpretation can be justified in terms of the theory. Otherwise he will be testing the validity of his own reinterpretation of the theory rather than the theory itself. This is not to exclude the legitimacy of redesigning tasks on the basis of conscious reinterpretation of the theory. Later in this article a number of elements essential to valid testing for conservation will be derived from an analysis of conservation tasks in terms of Piagetian theory.

Horizontal Decalage

The phenomenon of horizontal decalage complicates all four kinds of research activity, but especially those of verifying sequences and synchronisms, and setting up scales of cognitive growth. Simply put, the problem is that different tasks, which may be equally valid tests for the same developmental level, intellectual ability, or thought structure, may not be equally difficult. Determining whether or not two given abilities are acquired in fixed order or synchronously obviously requires that the tasks used to infer these abilities are equally difficult. Flavell and Wohlwill (1969) have proposed a competence-performance model to meet this problem. The practical use of this model must be delayed however until the values of parameters representing task difficulty can be established empirically. The model also has the disadvantage that it requires the inference of an "ideal" competence which can never be measured directly. A practical course of action is to use a number of tasks for the same ability, together with the strategy of classifying some subjects as transitional cases. This course of action will be justified later in the article for the case of conservation.

Inference Criteria

The problem of inference criteria is that of deciding which aspects of the subject's behavior on the task constitute necessary and sufficient criteria for inferring the developmental level, ability, or thought structure for which the task is a valid test. If the criteria are necessary but not sufficient, type I errors will be made; while if they are sufficient but not necessary, type II errors will be made. Brainerd (1973) appears to have resolved the issue of whether judgments alone should constitute criteria, or whether judgments and explanations should be required. He showed that if care is taken in the design of the tasks, a correct judgment alone can provide a necessary and sufficient criterion, whereas the requirement of a satisfactory explanation constitutes a sufficient but not a necessary criterion. Braine (1970) also advocates the use of nonverbal tests. The approach taken in this article in the case of conservation testing will be to emphasize the necessity of the use of objective operational measures of the quantities in question, whether verbal or nonverbal tests are used.

Conservation Tasks

Before proceeding to analyse conservation tasks generally, the need for concern with these methodological problems will be further emphasized by

illustrating the wide variation in volume conservation testing methodologies which have been used by different researchers. Four specific examples will be described: procedures used by Inhelder (1968), Elkind (1961), Goodnow and Bethon (1966), and Phillips (1971).

In the Inhelder procedure, subjects were shown three different transformations of a clay ball. In each case they were questioned about the room occupied by the bodies in water, and were asked to predict water levels after immersion of the transformed object in water. Subjects were cross-questioned by the experimenter in a face-to-face situation. Classification of the subject was made by the experimenter on the basis of all of the subject's responses. This famous "clinical technique" has been well described by Flavell (1963).

In the Elkind procedure, subjects were shown only one transformation of a clay ball. Subjects were tested in groups and recorded their own responses to a series of prepared questions with pencil and paper. Subjects were asked for five responses: they were asked (a) to say whether the two balls occupied the same amount of room or space, (b) to predict whether the two objects would occupy the same amount of room or space should one be rolled out into a sausage, (c) to judge after the transformation was actually carried out whether the two configurations did occupy the same amount of room or space, (d) to explain their responses, and (e) to predict whether each configuration would cause equal rises in water level upon immersion in identical containers of water. Subjects were classified as conservers only if they gave satisfactory responses in every case.

In the Goodnow-Bethon procedure, subjects were shown only one transformation of a clay ball. They were shown the water level rise which occurred when the untransformed ball was immersed in a glass of water and were required to predict what the level would be should the final configuration be immersed in an identical glass of water. Subjects were then asked to explain their responses. They were interviewed individually with a prepared set of questions. Subjects were classified as conservers or non-conservers according to their predictions of the water level, and conservers were further classified according to their explanations of equality of levels: subjects who referred to weight in their explanations were included in one category and not in the other.

In the Phillips procedure, subjects were shown pictures of the initial configurations, the final configurations and the transformations. The initial and final configurations were different arrays of one-inch metal cubes, immersed in water or in a medium of small solid cubes. Subjects were questioned individually and required to judge whether the immersing medium would rise to the same level in its container for the two different arrays of cubes, or for the same array in two different positions in the container. Subjects were required to explain their answers and were classified as conservers only if both correct judgments and correct explanations were given.

Other procedures, different from each of these, could be cited, but these four suffice to illustrate the diversity among procedures which purport to test for volume conservation. The theoretical analysis which follows will show that each of these procedures is subject to one or more sources of removable systematic error.

Volume conservation may be simply defined. It is the recognition that a given mass of material which undergoes a change of shape (without change in density) does not change in the volume it occupies. The difficulty lies in devising a test which will allow the experimenter to decide unequivocally whether such recognition may be inferred to a given subject. Volume conservation tests are usually variations of a procedure first devised and used by Inhelder (Piaget & Inhelder, 1968). Some experimenters (e.g., Lovell & Ogilvie, 1961) have used procedures similar to ones described in *The Child's Conception of Geometry* (Piaget, Inhelder, & Szeminska, 1960). Both kinds of procedure are forms of various quantity conservation tests used by Piaget over the years to test for conservation of substance, length, weight, and so on. Conservation testing generally will now be subjected to a theoretical analysis.

There are three essential steps in an orthodox quantity conservation test. These are:

1. Two identical objects are shown to the subject in some initial configuration—e.g., two balls of clay.
2. One of the objects is transformed into a new configuration as the subject observes—e.g., one of the balls is rolled out into a sausage.
3. The subject is required to compare the transformed object with the untransformed object in terms of some operational measure of the quantity being tested—e.g., the subject is asked if the ball and sausage will weigh the same on a balance, or will raise the levels of water in two identical containers to the same height.

These three steps in a conservation test may be characterized further by four elements which are essential to task validity. The first element is *presence of both the final configuration and the initial configuration for comparison by the subject*. Elkind (1967) has shown the importance of starting in the first step with two objects, one of which remains unchanged throughout. Should only one object be shown, then a subject may appear to conserve merely by recognizing that it is the same object after the transformation as it was before. In other words the subject recognizes the identity of the two configurations. Mere identity, otherwise called "pseudo conservation" (Piaget, 1967) can occur when one essential aspect of the conservation problem is missing: immediate perceptual difference between the two configurations. When there is no initial configuration for comparison the subject has to imagine what the initial configuration was, and as Elkind (1967) puts it, through "memory falsification" the subject may arrive at pseudo conservation. Similarly, the final configuration must be present in the final comparison. Piaget (Piaget, 1967; Piaget & Inhelder, 1971) has demonstrated the existence in subjects of "false anticipatory images." Thus if a subject who recognizes identity of body in a transformation, his memory of the initial configuration or his anticipatory image of the final configuration may be such that the observed and imagined configurations seem identical.

The second element is *demonstration of the transformation*. Without the transformation, the test is not of conservation but of estimation of the physical quantity. According to Elkind (1967), estimation is a more difficult task than conservation. Moreover logical necessity, which is characteristic of conservation, is not possible for such an estimation. Phillips (1971) has

suggested that actual demonstration of the transformation may not be essential. He tested two groups of subjects for conservation, the transformations being demonstrated to subjects in one group, and described with the aid of pictures to subjects in the other group. Phillips concluded no significant difference ($p > .05$) between the proportions of conservers in the two groups. Nonetheless a close inspection of his results shows that higher proportions of conservers were found when actual objects were shown than when pictures were shown. In fact for all of the conservation tests combined (quantity, interior and exterior volume) the difference may be taken to be significant ($F = 2.60$, $df = 1,118$, $p < .20$). Thus to use pictorial presentation of the objects and transformations may result in type II errors.

The third essential element of a conservation test is *use of a variety of transformations*. Subjects who are in a stage of transition from nonconservation to conservation find some transformations more difficult than others. This problem comes under the heading of "horizontal decalage." Piaget (1971) has compared it with "friction," while the analogue used by Flavell and Wohlwill (1969) is "noise." These analogies imply the existence of an ideal or pure "motion" or "signal" which could be found in principle by creating measuring situations where the "friction" or "noise" is reduced to zero. Thus there is postulated an ideal or pure ability. Piaget's approach is to consider the problem to be in principle quite intractable, but not at all crucial—something of a nuisance. Flavell and Wohlwill quite correctly recognize the extreme difficulty the problem poses for the work of verifying the "sequential patterns of cognitive structure." They have proposed a competence-performance model, which involves an equation for predicting the probability of success on a given task by a given subject. The equation contains parameters representing both the "noise" and the "pure signal" (that is, the ideal or "true" competence of the subject). A fundamental objection to this model is that it involves a pure or ideal ability which can never be directly measured; but even if this is acceptable, the model cannot be put to practical use until the chore of generating empirical data to define its parameters is in hand. In the meantime, the precaution should be taken to use a range of transformations in conservation testing. Subjects who conserve on some transformations but not on others may then be classed as transitional cases. If a dichotomy of subjects is required, the transitional cases may be classed as conservers whenever type II errors are considered to be more costly, and as nonconservers whenever type I errors are considered to be more costly.

The fourth element is the *use of an objective operational measure for the quantity to be conserved*. The initial and final configurations are compared in terms of this objective, operational measure. Some examples of operational measures are: for length—distance to be walked, area—grass for cows to eat in a field, weight—balance with another object on an equal-arm balance, substance—amount to eat or drink.

The experimenter cannot be sure that the subject understands the question if it is asked with the use of terms as "same volume" or "just as heavy," because in their early stages of development, quantity concepts are gross or undifferentiated (*global*). Verizzo (1970) found that terms such as "mass," "weight," and "volume" were virtually synonymous even for very bright adolescents. Furthermore, the experimenter should be careful not to

induce the subject to think only in perceptual terms of the quantity being tested. In Piagetian theory, the achievement of conservation of a physical quantity means that the subject has progressed from (a) notions of the quantity in terms of the egocentric or perception-based effects by means of which he came to be aware of it, to (b) understanding of the quantity in terms of objective interaction between bodies (Piaget & Inhelder, 1968, p. 59). Some experimenters ask comparison questions such as: "Now is the sausage heavier than the ball, or lighter, or the same?" without reference to a weighing balance. To do this is to lead the subject to rely on egocentric notions of the weight (force on the hand or "hardness"). For instance when Hyde (1970) tested weight conservation with the ball-sausage transformation, she actually asked subjects to heft the ball and sausage in their hands, thus inducing them to answer in terms of their perceptual judgments. Some of Hyde's nonconservers of weight might have conserved if the question had been posed using a balance scale. In addition, according to Griffiths, Shantz, & Sigel (1967) there may be problems associated with the use of relational terms, such as "more," "less," and especially "same," at least for preschool children. Some experimenters have resorted to "nonverbal" methods (e.g., Braine, 1968; Mermelstein & Shulman, 1967). In fact the experimenter who devises nonverbal methods is forced to use objective operational measures for the quantities in the conservation tests. Apparently, the use of nonverbal methods has indicated that reliance on the verbal mode in conservation testing may increase the probability of type II error. The preferred course of action is to use nonverbal methods whenever these can be devised and administered readily, but the nonverbal tests usually require either training sessions or complicated apparatus or both. Whenever verbal methods are more feasible, the experimenter must remember that type II errors may be made, and must ensure that the quantity to be conserved is referred to in terms of an objective operational measure, avoiding such words as "longer," or "smaller," or "weighs the same."

To recapitulate, there are four elements in the three-step conservation test which are essential to task validity:

1. presence of both the initial and final configurations,
2. demonstration or at least description of the transformation,
3. use of variety of transformations, and
4. use in the final comparison of an objective, operational measure of the physical quantity to be conserved.

The first and second of these elements place constraints upon the testing procedure. The third and fourth allow for some variation in procedure in terms of the transformations used and the specific operational measure of the physical quantity employed. The possible range of variations is limited by the requirement to avoid reinterpreting the theory. Thus specific design of a conservation task involves selection of the transformation to be used and selection of an objective operational measure of the quantity to be conserved. This will now be done for the case of volume conservation, starting with transformations.

Transformations

Piaget and his collaborators have used three kinds of transformation in volume conservation tests: rearrangement of arrays of solid cubes, deforma-

tion of spheres of clay (plasticene), and the dissolving of sugar in water. In the first method (Piaget et al., 1960), the subject has to recognize that the volume or room occupied by a stack of cubes (say $4 \times 6 \times 10$) does not change when the cubes are restacked in a different way (say $6 \times 5 \times 8$). In the second method (Piaget & Inhelder, 1968; Inhelder, 1968), a sphere of clay is either (a) rolled out into a "sausage," (b) flattened out into a "pancake," or (c) broken into several pieces. In the third method (Piaget & Inhelder, 1968; Inhelder, 1968), subjects are shown sugar lumps being placed in water and subsequently dissolving. Subjects note the rise in water level on addition of the lumps, and are then required to say whether the increased level will be maintained as the sugar dissolves.

According to Piaget, "There is a good deal to be gained in avoiding the use of discontinuous elements . . . with older children who would otherwise be able to solve the problem by arithmetical correspondence without considering the volume as such" (Piaget et al., 1960, p. 356). This means that the restacking of cubes method may be prone to type I errors, especially with older children. On the other hand there is the chance of type II errors if a very difficult or unusual transformation is used, because to some extent a volume conservation test involves understanding the transformation. To understand the transformation requires the construction of a grouping of physical or infralogical operations to represent the transformation (Piaget & Inhelder, 1968). Such groupings involve "atomistic schemes," in which shape changes are represented by movements of elementary pieces of the object from one place to another in the mass (these movements being the physical operations). There are physical transformations in which volume is not conserved (e.g., freezing, or mixing of some liquids), and obviously mature volume conservers are readily able to assimilate such transformations. Therefore two achievements may underlie conservation of volume in a given transformation: (a) volume conservation as defined above, which involves differentiation of volume of a body from other properties (shape, surface area, etc.) and (b) understanding of the transformation, which involves construction of a grouping of physical operations. Type II errors may occur if the subject has the first ability (volume conservation per se) and not the second (understanding of the transformation). Consequently very difficult transformations should be avoided. To understand the deformation of clay requires no more than an atomistic scheme for envisaging the transfer of parts of the clay from one location in the mass to another. This kind of scheme is very similar to the scheme of additive compensations, which is acquired in the concrete-operational period. Volume conservation however requires the scheme of multiplicative compensations, which is acquired in the formal-operational period (Inhelder & Piaget, 1958). This means that for deformation of clay, the kind of type II error just described is low since the understanding of the transformation requires an atomistic scheme which is accessible to thinking at an earlier developmental level than at least one scheme (multiplicative compensations) necessary for volume conservation. On the other hand, the atomistic scheme necessary to understand dissolution of sugar involves the separation from one another into the water of invisible particles of sugar, *each with a fixed volume* and such that the total volume of the sugar lump is equal to the sum of the volumes of these individual particles. Therefore,

this scheme itself presupposes volume conservation, so that the kind of type II error just described may be more likely for this transformation than for deformation of clay. Furthermore, there is in fact a contraction of volume as the sugar dissolves (even though very small) which might make the test suspect for older children who might be aware of this (from science classes perhaps).

Objective Operational Measures

Careful investigation into the relative difficulties of various transformations may provide better understanding of the development of volume conservation. Uzigiris (1964) found that difficulty varied as a function of stimulus material, while an experiment by the writer (1973) indicated that the relative magnitude of the change in surface area may be the important factor determining difficulty. In the absence of any firm evidence about the relative difficulty of transformations, and in view of the preceding analysis, a prudent course would be to employ the three deformation of clay transformations or minor variations of these.

Consider now the operational measure to be used in the comparison of volumes of the initial and final configurations. Most experimenters who have employed an operational measure for volume have, like Piaget and Inhelder, used the rise in level which occurs when an object is placed in a container of water. There is a problem with this measure in that the subject must not only conserve volume, but also realize that the volume and not some property of the body determines the amount of change in the water level. Piaget's research has shown that in the concrete-operational period children believe the rise in the level to be caused by the object's weight (Piaget & Inhelder, 1968). Therefore if volume conservation is tested by asking only if the initial and final configurations would cause the same rise in water level, a subject may predict equality of the rises in level on the basis of conservation of weight rather than of volume. An experiment by the writer (1973) showed that up to 12% of junior high school children may appear to conserve volume by conserving weight when the test involves questions about water levels only, without references to volume or room occupied. This represents about 16% of those who would be classed as conservers on such a test.

In Vinh-Bang and Inhelder's experiments, from 10% to 30% of subjects (depending on the age group) explained equality of water levels in the volume conservation test by action of weight (Piaget & Inhelder, 1968, p. xiii; Inhelder, 1968, p. 317). The reader must decide whether these subjects should be classed as conservers. Goodnow and Bethon (1966) gave two sets of results for volume conservation acquisition; one including all subjects who anticipated equality of water level, the other excluding subjects who explained this equality by reference to equal weight. Of a group of "average ability" eleven-year-old boys, 56% were in the first category, 31% in the second.

To be so ambiguous about the results of a volume conservation test is both undesirable and unnecessary. Subjects who correctly predict equality of water levels merely by conserving weight should not be classed as conservers of volume. To avoid the possibility of type II errors which might then occur, efforts should be made to ensure that the subject understands

the displacement of water level to be a measure of space or room occupied by the body immersed. In other words, steps should be taken to make sure subjects are aware that the question being asked refers to the volume or room occupied by the two bodies. When this precaution is taken, subjects who persist in explaining equality of water levels in terms of weight should not be classed as volume conservers.

In fact care was taken in the original volume conservation experiments to direct subjects' attention to the volume occupied by the bodies in the test. Piaget and Inhelder describe the procedure as follows:

We decided after trying various methods to have volume estimated by means of the space occupied by the ball and by its successive forms in a glass of water, this space being identified by the level of water. We mark the initial level of water with ink or with an elastic band. Then we ask, "If I put this ball in the water, will it take up room? Will the water rise or stay at the same level?" This first question is suggestive by design. Indeed we have established elsewhere that a certain proportion of subjects between five and eight years are ignorant of the phenomenon and are surprised to see the water rise upon immersion of the solid. On the other hand, we have been able to establish that most subjects before eight or nine years explain the rise in level by referring not so much to the volume as to the weight, which produces according to the child an upwards current. Now these two reactions are very interesting in the study of conservation of volume. . . .

But the matter at hand is to set the problem from the start in the field of volume. That is why we say to the child "take up room." [*prendre de la place*] After having established in the subject that the first ball made the level rise and having marked the second level, we present a second ball, similar to the first, asking how far it will bring up the level. Every child referred to in this chapter indicated the same second level and thereby showed himself capable of understanding the basic ideas in the problem. It is then and only then that the second ball is transformed into a sausage, disc, etc., or broken into pieces, with the question each time: "and now does it take up the same room in the water as the other ball? Where will the water rise to?" etc. (Piaget & Inhelder, 1968; writer's translation).

A further control is described by Piaget (Piaget & Inhelder, 1968, Introduction, pp. xvii-xviii): before comparison of the sausage and ball by the immersion in water, a metal ball is substituted for the original untransformed clay ball. Subjects who conserve weight but not volume would give an incorrect response in this test.

Now as careful as these precautions are, justification can be advanced for going beyond them. The last control mentioned is necessary to avoid classifying subjects as volume conservers when in fact they have done no more than conserve weight in the test. This control may however act as a "trap" for volume conservers who persist in explaining rise in water level in terms of the weight of the immersed object. To avoid this, the subject should be shown by explicit demonstration that (a) objects of equal volume but different weight cause equal rises in water level, and that (b) objects of equal weight but different volume cause different rises in water level. An objection to this procedure might be that volume conservation may be induced in subjects who would have been classified as nonconservers before the demonstration. Indeed, Vinh-Bang and Inhelder have shown that learning of volume conservation can occur rapidly in some subjects when they are shown immersion of ball sausage etc. in water. (Percentage of conservers went from 37% to 78% of eleven-year-olds after three

demonstrations—Inhelder, 1968, p. 318; Piaget & Inhelder, 1968, Introduction p. xviii.) An experiment by the writer (1973) indicated that discovery of the fact that the volume of the immersed body determines the rise in water level may precipitate acquisition of volume conservation in older subjects.

Inference Criteria

A reply to this objection necessitates consideration of the problem of inference criteria. Success or failure on a volume conservation test infers presence or absence of certain schemas, (e.g., proportionality and multiplicative compensations) and hence, at least the beginning of formal operations. To confine attention to what a subject actually *does* in a given test rather than to what he may be *able* to do under appropriate circumstances would be to allow temporary aberrations to be confused with absence of the theoretical underlying abilities. Consider a given subject who is unaware that an object's volume determines displacement of water, even though he has begun to construct formal operations. He conserves weight and is in possession of schemes of multiplicative compensations and proportionality, but yet may fail in a given test to conserve volume. Such failure would be merely a temporary mistake; such a subject would be found to acquire volume conservation very quickly in the face of a few simple demonstrations. To classify such a subject as a nonconserver would be to commit a type II error.

In the Piagetian clinical technique, the subject is given every chance to arrive at the correct answer. A temporary mistake (or for that matter a chance correct response) is not accepted at face value, but the subject is questioned further (Inhelder, 1968, pp. 344-346). Of course the clinical technique has its problems, associated with the subjectivity of classification criteria. A more objective, standardized procedure is sought here; one which can be administered to individual subjects or, if necessary, to groups of subjects recording their own responses with pencil and paper. In such a test, two kinds of response may be recorded: a subject's *judgment* of the equality of the two volumes, and his *explanation* of the equality. Brainerd (1973) argues that the thought structures to be inferred from conservation tests are, according to Piaget, relatively independent of language. Therefore adequate explanations are sufficient but not necessary conditions for the inference of a given thought structure. In addition (he argues) to use explanations as necessary criteria is to confine the response mode to the verbal, and this "unduly restricts the behavioral domain to which the theoretical construct (structure) applies." Thus the use of explanations as the sole criterion is liable on two counts to result in failure to infer the existence of a thought structure that does exist, i.e., explanations criteria "are subject to at least two sources of systematic type II error." Judgments, on the other hand, "from the standpoint of Piagetian theory . . . are not subject to any known sources of systematic error."

A Prototype Volume Conservation Test

Taking into account the requirements derived above, a test for volume conservation was designed and administered to a grade five class in an Edmonton (Canada) school. The test has three special features. First, a preparation of subjects was undertaken so as to ensure that they did

associate the objective operational measure—displacement of water level upon immersion of an object—with occupied volume. Second, the test avoided the use of relational terms such as “more,” “less,” or “same.” Third, a nonverbal response was required. An added but nonessential feature was the administration of the test by means of a videotape recording. Use of the video medium allowed the test to be administered to a whole class at one time, and avoided unwanted or uncontrolled subject-experimenter interactions. The videotape presentation was supplemented by actual “live” demonstration of the transformations. This avoided problems due to the distortion which accompanies the two-dimensional electronically-mediated TV picture.

The test will now be described as it would be administered to an individual subject by a “live” experimenter. Parenthetical notes will be used to describe those adaptations which were made for the videotape presentation.

1. *E* presents three identical glasses to *S*. The glasses are side by side and filled to the same level with water. The levels are marked by a black line on each glass. *E* points out that the levels are the same in each glass. The first glass remains unchanged throughout the experiment.

2. *E* displays five or six identical balls of plasticene and notes that they are all the same. He asks “What do you think will happen if I put this ball into this [second] glass? Will it take up any room in the water? Where will the water come to?”

3. When *S* has responded, *E* says “let’s try it,” and places the ball into the second glass. *E* says “See? The water came up to here.”

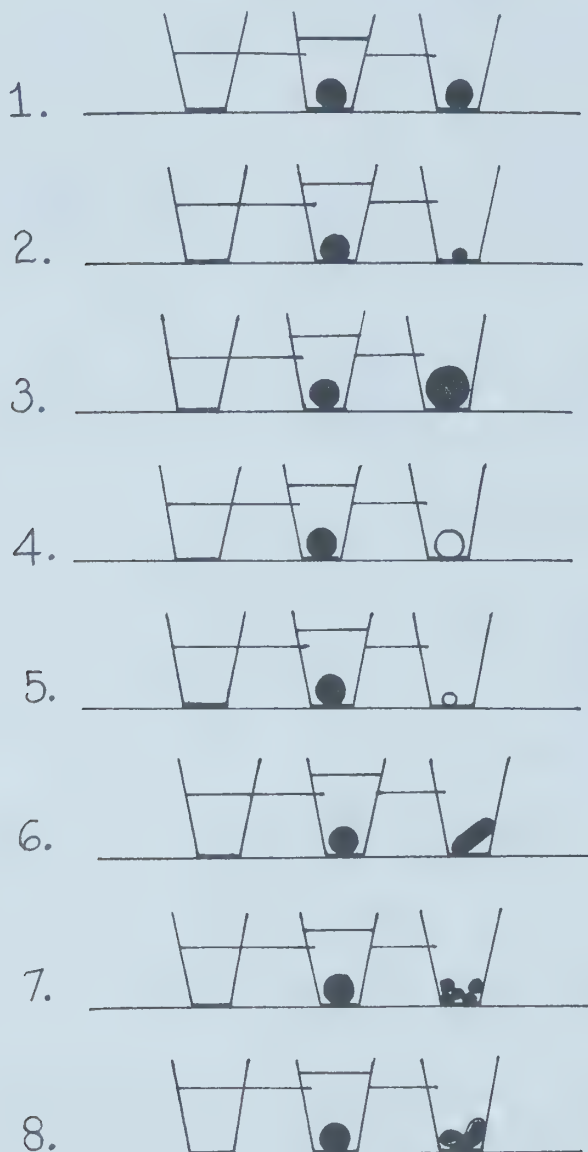
4. *E* takes another ball identical to the first and says “Now what if I put this ball into this [third] glass? How much room will it take up in the water? Where will the water come to? Show me.” *S* is required to respond by pointing to where he thinks the level will rise to in the third glass upon immersion of the ball. (In the videotape presentation, subjects were required to respond by marking in a level on a diagrammatic representation of the three glasses side by side, with the levels shown only for the first two. The answer sheet on which subjects responded is shown in Figure 1.) Note that the response is nonverbal in both cases, and that relational terms are not used in posing the question.

5. After *S* responds, *E* says “Let’s try it,” places the ball in the third glass, and says “See? The water came to here.”

6. Steps 4 and 5 are repeated using respectively (a) a larger plasticene ball, (b) a steel ball of the same size, and (c) a smaller lead ball. Each time the same question is asked: How much room will *this* ball take up in the water. If I put this ball into the glass, where will the water come to?” Each time, after *S* responds, the ball is placed into the third glass to show how much the water level rises. (In the videotape presentation, a fresh prefilled glass was used each time as the third glass.)

The procedure to this point represents the preparation of the subject. By responding incorrectly to the last question in the preparation, subjects indicate that they are persisting in associating the displacement of the water level with the weight of the immersed object. Such subjects should be classed as nonconservers to avoid type I errors, for these subjects may

FIGURE 1
RESPONSE SHEET FOR VOLUME CONSERVATION TEST



appear to conserve volume by conserving weight. On the other hand if type II errors are more costly, the preparation procedure may be repeated before classifying such subjects as nonconservers. With the preparation complete, *E* proceeds to the actual volume conservation tasks.

7. *E* presents two balls identical with the one immersed in the second glass. One is deformed into a cylinder or "sausage." *E* says "How much room will the sausage take up in the water? If I put the sausage in the [third] glass, where will the water come to?" *S* responds, but this time *E* does not actually immerse the transformed object. (In the videotape presentation, the transformation and comparison with the untransformed ball were demonstrated "live," interrupting the playback of the tape.)

8. Step 7 is repeated for two other transformations: (a) breaking the ball into four or five pieces, and (b) flattening the ball into a disc or "pancake."

Subjects who are successful in the preparation and respond correctly to all three volume conservation tasks are classified as conservers. Subjects

who are successful in the preparation and respond correctly to only one or two of the volume conservation tasks are classified as transitional cases. Subjects who are unsuccessful in the preparation or who respond incorrectly to all of the volume conservation tasks are classified as nonconservers. Success in the preparation is measured by response to the last question. (An objective method of judging whether the lines drawn by *S* were correct was used in the test run: the line drawn by *S* had to come within 0.5 mm of the correct level at some point to be judged correct. This was determined by centering a 1 mm thick line drawn on a transparent sheet over the correct level.)

In the grade five class tested ($N = 28$), 14 subjects were classified as conservers, 10 as nonconservers, and 4 as transitional cases. Only one subject responded incorrectly to all of the preparation questions and was thereby classified among the nonconservers. The average age of the subjects was 10 years 10 months. No problems arose in the administration of the test or in its scoring, and subjects indicated to the experimenter that they understood what they were required to do. The entire testing procedure took just over 10 minutes.

Summary

This volume conservation task provides an illustration of how methodological difficulties associated with Piagetian tasks can be overcome by means of careful design of the testing procedure based on analysis of the tasks. The methodological difficulties defined were concerned with task validity, horizontal decalage, and inference criteria. Consideration of validity led to the conclusion that design of the task has to be justifiable in terms of Piagetian theory, unless conscious reinterpretation of the task is intended. In conservation tasks, this means that the transformation should be demonstrated, and that the initial and final configuration have to be simultaneously seen and compared by the subject. Consideration of horizontal decalage led to the conclusion that a variety of forms of the task should be employed. In conservation tasks this means that a number of transformations should be demonstrated, and that very difficult transformations should be avoided. Consideration of inference criteria led to the conclusions that nonverbal methods should be used wherever feasible, that judgments alone represent sufficient criteria, and that in conservation tasks, objective operational measures should always be used for the quantity to be conserved.

Assistance in preparation of this article was provided by Dr. C. Brainerd and Dr. L. Stewin of the University of Alberta.

References

- Beilin, H. The training and acquisition of logical operations. In M. F. Rosskopf, L. P. Steffe, & S. Taback (Eds.), *Piagetian cognitive-developmental research and mathematical education*. Washington: National Council of Teachers of Mathematics, 1971.
- Braine, M.D.S. The ontogeny of certain logical operations: Piaget's formulation examined by nonverbal methods. In I.E. Sigel & F.H. Hooper (Eds.), *Logical thinking in children*. New York: Holt, Rinehart & Winston, 1968.

- Braine, M. D. S. Piaget on reasoning: A methodological critique and alternative proposals. In Society for Research in Child Development. *Cognitive development in children—Five monographs of the Society for Research in Child Development*. Chicago: The University of Chicago Press, 1970.
- Brainerd, C. J. Continuity and discontinuity hypotheses in studies of conservation. *Developmental Psychology*, 1970, 3, 225-228.
- Brainerd, C. J. Judgements and explanations as criteria for the presence of cognitive structures. *Psychological Bulletin*, 1973, 79, 172-179.
- Elkind, D. Quantity concepts in junior and senior high school students. *Child Development*, 1961, 32, 551-560.
- Elkind, D. Piaget's conservation problems. *Child Development*, 1967, 38, 15-27.
- Flavell, J. H. *The developmental psychology of Jean Piaget*. New Jersey: D. Van Nostrand, 1963.
- Flavell, J. H., & Wohlwill, J. Formal and functional aspects of cognitive development. In D. Elkin & J. H. Flavell (Eds.), *Studies in cognitive development: Essays in honor of Jean Piaget*. New York: Oxford University Press, 1969.
- Goodnow, J. J., & Bethon, G. Piaget's tasks: The effects of schooling and intelligence. *Child Development*, 1966, 37, 537-582.
- Graves, A. J. Attainment of conservation of mass, weight, and volume in minimally educated adults. *Developmental Psychology*, 1972, 7, 223.
- Griffiths, J. A., Shantz, C. A., & Sigel, I. E. A methodological problem in conservation studies: The use of relational terms. *Child Development*, 1967, 38, 841-848.
- Higgins-Trenk, A., & Gaite, A. J. H. Elusiveness of formal operational thought in adolescents. *Proceedings of the 79th Annual Convention of the American Psychological Association*, 1971, 6, 201-202.
- Hobbs, E. D. Adolescents' concepts of physical quantity. *Developmental Psychology*, 1973, 9, 431.
- Hyde, D. M. G. *Piaget and conceptual development*. London: Holt, Rinehart & Winston, 1970.
- Inhelder, B. *The diagnosis of reasoning in the mentally retarded*. New York: John Day, 1968. Trans. from the French *Le diagnostic du raisonnement chez les débiles mentaux*. (2nd ed.) Neuchatel: Delachaux et Niestle, 1963. (1st ed. 1943).
- Inhelder, B. N., & Piaget, J. *The growth of logical thinking from childhood to adolescence*. New York: Basic Books, 1958. Trans. from the French *De la logique de l'enfant à la logique de l'adolescent*. Paris: Presses Universitaires de France, 1955.
- Keller, H. R., & Hunter, M. L. Task differences on conservation and transitivity problems. *Journal of Experimental Child Psychology*, 1973, 15, 287-301.
- Lefrancois, G. R. A treatment hierarchy for the acceleration of conservation of substance. *Canadian Journal of Psychology*, 1968, 22, 277-284.
- Lovell, K., & Ogilvie, E. The growth of the concept of volume in junior school children. *Journal of Child Psychology and Psychiatry*, 1961, 1, 191-202.
- McLaughlin, J. A., Bentler, P. M., & Stephens, B. Monotonicity analysis of Piagetian assessments of reasoning. *Proceedings of the 78th Annual Convention of the American Psychological Association*, 1970, 5, 283-284.
- Mermelstein, E., & Shulman, L. S. Lack of formal schooling and the acquisition of conservation. *Child Development*, 1967, 38, 39-52.
- Phillips, D. J. The development of the concept of displacement volume: A hierarchical model and its partial testing under two methods of presentation. *Journal of Research in Science Teaching*, 1971, 8, 9-19.
- Piaget, J. Untitled article. *Contemporary Psychology*, 1967, 12, 532-533.
- Piaget, J. The theory of stages in cognitive development. In D. R. Green, M. P. Ford, & G. B. Flamer, (Eds.) *Measurement and Piaget*. New York: McGraw-Hill, 1971.
- Piaget, J., & Inhelder, B. *Le développement des quantités physiques chez l'enfant*. (3rd ed.) Neuchatel: Delachaux et Niestle, 1968. (1st ed. 1941).

- Piaget, J., & Inhelder, B. *Mental imagery in the child*. London: Routledge & Kegan, Paul, 1971. Trans. from the French *L'image mentale chez l'enfant*. Paris: P.U.F., 1966.
- Piaget, J., Inhelder, B., & Szeminska, A. The child's conceptions of geometry. London: Routledge & Kegan, Paul, 1960. Trans. from the French *La géométrie spontanée de l'enfant*, 1947.
- Sigel, I. E., & Hooper, F. H. *Logical thinking in children*. New York: Holt, Rinehart & Winston, 1968.
- Somerville, S. C. The pendulum problem: Patterns of performance defining developmental stages. *British Journal of Educational Psychology*, 1974, 44, 266-281.
- Uzigiris, I. C. Situational generality of conservation. *Child Development*, 1964, 35, 831-841.
- Verizzo, O. Conceptions of conservation and reversibility in children of very superior intelligence. *School Science and Mathematics*, 1970, 70, 31-36.

S R . R U T H E A G A N

The University of New Brunswick

An Investigation Into the Relationship of the Pausing Phenomena In Oral Reading and Reading Comprehension

This study attempted to investigate whether misuse of the pause in oral reading was related to the silent and oral reading comprehension abilities of young children.

Although the pausing phenomena may be considered as one product of the reading act, in this study it was considered to be an overt manifestation of how children might be processing visual input while reading silently and orally. The results of the study revealed that the pausing phenomena in oral reading were consistently and significantly able to discriminate between silent reading ability groups.

However, while the "processes" of silent and oral reading seemed to be similar, the "products" (or comprehension) of silent and oral reading did not appear to be related. Further, those children who were the better silent readers seemed to be impeded by the necessity of reading orally. Nevertheless, oral reading appeared to assist comprehension for the younger and less proficient readers. (Sr. Eagan is Associate Professor in the Faculty of Education. The University of New Brunswick.)

This study was an attempt to determine the nature of the pausing phenomena in oral reading, and whether these phenomena are possible overt features of the verbal processing entailed in reading orally, and also in reading silently, by relating them to the silent and oral reading comprehension abilities of the subjects.

In this study, the visual input, or orthographic string, was not organized for the child. It was assumed that how the child used the pause in his oral reading indicated how he was organizing the visual input. By investigating the number, length and placement of pauses as the child used these devices in oral reading, it was hoped that cues would be obtained to indicate whether the child was organizing the reading material into meaningful word groupings, or whether by pausing within such groups, he was disrupting the linguistic and cognitive unit.

The unique feature of this study was that the method by which the pausing phenomena were measured was a completely objective and repeatable procedure.

Purpose of the Study

The purpose of the study was to investigate the relationship between pauses made by children reading orally, and their reading comprehension ability, both silent and oral, and in particular, to determine whether the use, or rather misuse of the pause (by disrupting a linguistic unit) was related to silent reading comprehension groups, and oral reading comprehension scores.

The study also attempted to test the assumption that young children learning to read use oral reading processes that are similar to those they use when reading silently.

Definition of Terms

The following are some of the key definitions used in this study:

Pause: in this study, the term "pause" referred to the interruption or termination of the voice stream in oral reading. It was measured by an *Esterline Angus Speed Servo* machine, in milliseconds (See Appendix A).

Criterion Pause: since pause duration shows marked variability across oral readers, a "criterion pause" was established for each subject in the study. This was accomplished by calculating the mean length of the between-word pauses produced by subjects in their oral reading of a set of syntactic constituents (defined below).

Significant Pause: a significant pause exceeded the duration of an individual's "criterion pause" by a critical amount. This critical interval was precisely determined by means of pilot work. For purpose of this study, a significant pause was 2.5 times longer than the criterion pause.

Syntactic Constituent: this referred to the lowest major constituent as defined by Latham (1972). It may be formed from all labelled nodes in a surface structure tree. These constituents are found by locating those nodes which are immediately above the lexical nodes and deciding whether or not there are sister-nodes to the lexical nodes. If there are no sister-nodes to a specific lexical node, then the node immediately dominating that lexical node is a lowest major constituent. If there are sister-nodes to a specific lexical node, then the lowest major constituent associated with that node is the node which immediately dominates all sisters of the lexical node in question (See Appendix B).

Word Recognition: referred to the ability of the child to look at a word and pronounce it aloud correctly, as determined by the judgment of the investigator.

Theoretical Framework

Pause can serve two obvious purposes in the processing of language: (1) it serves to separate the string into meaningful units for analyzing content (Osgood, 1954; Garrett, Fodor, & Bever, 1966; Neisser, 1967;

VanUden, 1970; Bond, 1971) and (2) it facilitates the storage of items in short-term memory (Dale, 1964; Miller & Isard, 1963; Wickelgren, 1965, 1969; Neisser, 1969; O'Connell, Turner and Onoska, 1968; Smith, 1971). Any disturbance in these two uses of the pause in the processing of written language by children learning to read, may seriously impair their comprehension.

Research with adults has indicated that the grouping of words in reading orally is very consistent (North & Jenkins, 1951; Kolers, 1970; Cromer, 1970; Wilkes & Kennedy, 1970; Martin, Kolodziej, & Genay, 1971; Goldman-Eisler, 1968). Pause in oral reading by adult fluent readers always occurred at grammatical boundaries. Since a group of words, separated by pauses, has a certain cohesiveness, it follows that if this group is meaningful, it can more readily be stored as a unit in the short-term memory. It also follows that if this group lacks cohesiveness because of pausing within the group itself, then comprehension is likely to deteriorate.

Since the major goal of reading is comprehension, it is reasonable to suppose that the perceptual and cognitive processing of the material would be facilitated if the units stored for processing corresponded to the major sense groupings in the material being read. It seems then, that if the words of a syntactic constituent group were stored in the short-term memory as one unit, that comprehension would be facilitated. If, on the other hand, pausing occurs within the syntactic constituent group, it could be assumed that comprehension would be adversely affected.

Reading begins with the graphic input and ends with meaning as output. What happens between input and output has been hypothesized by many researchers, and is often called the reading "process". Since one cannot get inside a child's head to find out what is happening there, one way of attempting to investigate the reading process is to examine the output, or the "product". This product can be analyzed in two ways: (1) by examining the child's comprehension of what he has read, and (2) by analyzing the oral output in an oral reading performance. A child learning to read is, for the first time in his linguistic experience, exposed to the syntactic structure of the language from which the prosodic structure (one feature of which is pause) is absent. The present study is concerned with the use of the pause as a behavioral manifestation of how the young reader may be using the surface structure relationships to comprehend the message.

The models of beginning silent reading (Figure 1) and beginning oral reading (Figure 2) both indicate that the child recodes to aural input and/or oral output before he decodes the meaning, or comprehension. If during this intermittent step between graphic input and meaning the young beginning reader loses the cohesiveness of the word group because he pauses within this group, then such a disjunction may interfere with comprehension.

FIGURE 1
AN ADAPTATION OF GOODMAN'S 1970 MODEL
TO ILLUSTRATE EARLY SILENT READING

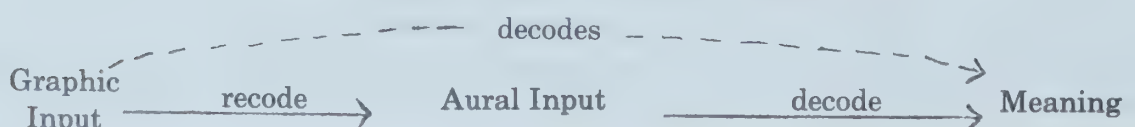
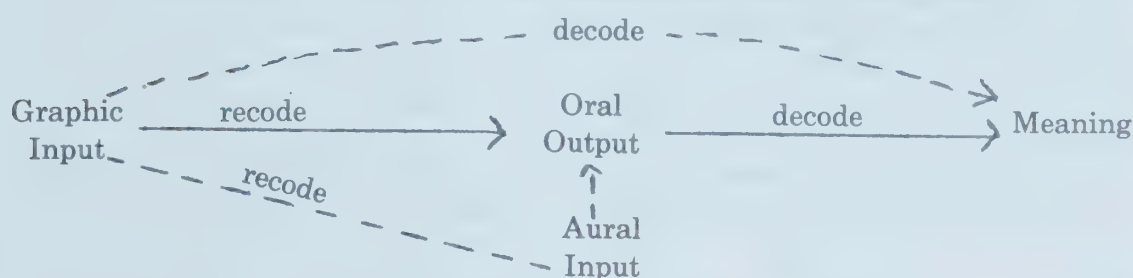


FIGURE 2
INVESTIGATOR'S MODEL OF EARLY ORAL READING



If, as assumed in this study, the silent and oral reading processes of young beginning readers are comparable (see Figures 1 and 2), then an analysis of how these young children use timing cues (pauses) in their oral reading, and the relationship of the use of these cues to comprehension (whether silent or oral comprehension), may provide some insights into some of the learning to read processes.

Hypotheses

Following are the main null hypotheses tested by this investigation:

1. That average readers, above-average readers and below-average readers do not differ significantly in the percentage of total reading time spent in pausing while reading orally,
 - a. when the effect of inadequate word recognition ability is eliminated from the data;
 - b. when the effect of inadequate word recognition ability is included in the data.
2. That average readers, above-average readers and below-average readers do not differ significantly in the period of time spent pausing within syntactic constituents while reading orally,
 - a. when the effect of inadequate word recognition ability is eliminated from the data;
 - b. when the effect of inadequate word recognition ability is included in the data.
3. That average readers, above-average readers and below-average readers do not differ significantly in the number of pauses made within syntactic constituents while reading orally,
 - a. when the effect of inadequate word recognition ability is eliminated from the data;
 - b. when the effect of inadequate word recognition ability is included in the data.
4. That there is no significant relationship between (1) the percentage of pause time used by children reading orally, (2) the period of time spent pausing within syntactic constituents while reading orally, (3) the number of pauses made within syntactic constituents by children reading orally, and
 - a. their oral reading comprehension scores.
5. That there is no significant relationship between (1) the percentage of total reading time spent in pausing, (2) the period of time spent in pausing within syntactic constituents, (3) the number of pauses made within syntactic constituents, and
 - a. Auditory Memory Span for Digits Forward

- b. Auditory Memory Span for Digits Backward
- c. Visual Memory Span for Letters, and
- d. Intelligence Quotients.

The Experimental Design

Sample

A stratified random sampling procedure was utilized to select the sample. Twelve average, twelve above-average, and twelve below-average readers were selected from each of grades two and three, from four schools, making a total of 72 children—six cells consisting of 12 children each. In each cell the sexes were equally divided.

Testing Instruments

Silent reading comprehension ability was measured by the *Gates-MacGinitie Silent Reading Test*. Oral reading performance was determined by the *Gilmore Oral Reading Test Form C* (1968). The *Lorge-Thorndike Intelligence Test, Level 2, Form A* was administered to all the subjects. The auditory memory span subtests of the *Weschler Intelligence Scale for Children* (1949), and the visual memory span subtest of the *Detroit Test of Learning Aptitudes* (1967) were used to measure memory span. A Word Recognition Test was composed by the investigator and included all the words occurring at least once in the *Gilmore Oral Reading Test, Form C*.

The investigator's copy of the *Gilmore Oral Reading Test, Form C* was divided into syntactic constituents, using the linguistic algorithm proposed by Latham (1972) to determine syntactic constituents (see Appendix B). The *Esterline Angus Speedservo AZAZ portable Labgraph* recorded on a strip chart, a graphic display of the speech sounds, as oscillations to the right of the baseline on the ten-point amplitude scale of the chart, while pauses were displayed as time intervals along the baseline (see Appendix A).

Statistical Analysis

The basic design of the study was a three-way Analysis of Variance over silent reading comprehension groups, on a number of pause measurements obtained from the oral reading performance of the test sample. These pause measurements constituted the main dependent variables in the study and were obtained objectively. Other dependent variables included auditory and visual memory span measurements, and intelligence.

In addition, the relationship between the oral reading comprehension scores of the test sample, and the various dependent variables included in the study, were examined by means of correlation coefficients and a three-way Analysis of Variance, grouping by oral reading comprehension, grade and sex.

Data were analyzed over the entire oral reading material read by each subject, and also over the first 70 syntactic constituents, which constituted the test material read in common by the entire sample.

Main Findings and Conclusions

The percentage of reading time spent pausing while reading orally was able to discriminate between silent reading comprehension groups of varying degrees of ability. This finding tends to lend support to the contention

of Pival (1971), Goodman (1968) and McCracken (1967) that the silent and oral reading processes of young children may be very similar.

Although the time spent pausing while reading orally was able to discriminate significantly between silent reading comprehension groups, all the children tended to resist a disruption within a syntactic constituent if at all possible. This finding agrees with the research on the processing of oral language, especially that of Wilkes and Kennedy (1970), Martin, Kolodziej, and Genay (1971), and Brown (1971), even though their research was done with adult readers.

The number of pauses the subjects made within syntactic constituents, however, seemed to depend more on the semantic and syntactic difficulty of the reading material than on their word recognition abilities. This finding seemed to support Goodman and Niles (1970) statement that the perceptual, syntactic and semantic information used in the reading process are used simultaneously and not sequentially; and also Brown's (1971) findings that 64 per cent of pause variance in oral reading can be predicted from syntactic measures.

Although all the pausing variables were able, in one way or another, to discriminate between silent reading comprehension scores, there was no statistical evidence in this study to indicate any relationship at all between oral reading comprehension scores and any of the pausing variables measured. Further, the data revealed no relationship between oral reading comprehension scores and silent reading comprehension scores.

Since the data do indicate that the use of the pause phenomena in oral reading does differentiate significantly between silent reading comprehension groups, it seems that the model of silent reading shown in Figure 1 appears to be correct, and that there actually may be an element of aural language input in silent reading which assists in comprehension, even when reading silently.

However, when the child is asked to read orally, it appears from these data that there is no relationship between how he is processing his visual input (i.e. syntactic constituents), and the meaning (or product) he obtains from the passages.

An examination of Figure 2 may provide two possible explanations for this occurrence: (1) in the process of recoding to aural input-oral output and then decoding to meaning, the child is unable to finish the process by recalling what he has "heard", either because he cannot remember or because he neglected to recode to aural input. He recoded directly to oral output from graphic input and did not "listen" to what he was reading; or (2) he stops the process at oral output and does not even attempt to complete processing for meaning. He is therefore only able to grasp what little comprehension of the material he has obtained from the element of direct decoding from graphic input to meaning, as shown by the broken line in the model. This could account for the lack of relationship between how the child seems to be processing his visual input while reading orally and oral reading comprehension, while at the same time the actual silent and oral reading processes may be very similar (see Figure 1 and Figure 2).

A perusal of the individual oral reading comprehension scores, however, did indicate that the younger and less proficient silent readers did seem to

be better able to "listen" to themselves while reading orally, probably because they are still using a strong auditory component while reading silently also. Consequently, the auditory "feedback" from an oral reading performance tended to help their comprehension of the passage.

There was a lack of significant relationship between memory span (both visual and auditory) and the processing of written language in this study. However, a closer inspection of the data suggested that the relationship between the processing of written language and memory span is something that develops, or is learned, but was not really evident in this study until towards the end of the grade three level.

Although intelligence did appear to have some relationship to how a child was using the pause while reading orally, the relationship was closely tied in with the age of the child, and the syntactic and semantic difficulty of the material which he was reading.

The data also indicated that semantic and syntactic difficulty of the reading material possibly has more bearing on the ability of the children to comprehend the material than does their ability to recognize the words.

Significant contributions to the effectiveness of reading, arranged in order of rank as determined by the data in this study seem to be: the efficiency of the reading process used by the child, the syntax and semantics of the material to be read, and lastly, the word recognition skills of the reader.

Implications of the Study

The implications of this study were divided into three sections: implications for (1) reading theory, (2) reading research, and (3) the teaching of reading.

Reading Theory

The data indicate that the syntactic constituent, as defined in this study, is very probably a functional linguistic and cognitive unit in the perception of written language.

Further, timing cues (or pauses) in oral reading, seem to be able to differentiate between the silent reading comprehension abilities of children who have been exposed to reading instruction for not more than 2.5 years, and also provide some indication of how these children may be processing written language.

The products of these processes (or the comprehension of the text), however, do not appear similar in the two situations—silent reading and oral reading. It seems then that although an observation of where pauses occur in oral reading output may be a good indication of the reading "processes" employed by young children, silent reading is a better instrument to use if the "product" (or comprehension) is to be observed.

There also seems to be a close relationship between the surface structure of the text, the number and length of pauses made while reading orally, and the silent reading comprehension ability of the reader. Pausing in oral reading might then be used to contribute to insights into whether a reader is able to simultaneously process the syntactic and semantic content of a passage, which may be the "linguistic sophistication" referred to by Chomsky and Halle (1968, p. 50) and which they contend is necessary for acceptable oral reading.

All the data reemphasized that the ability to process written language is a developmental process. At the grade three level there was sufficient evidence to indicate a very important developmental stage, and several factors occurred that were not present at the grade two level.

Reading Research

The pause (or juncture) is only one aspect of the entire suprasegmental system. Perhaps further research, based on this study and using the same, or similar instrument for obtaining objective measurements, could be initiated, using the suprasegmental phonemes of pitch and/or stress as the dependent variable(s).

An attempt might also be made to study the actual use of the pause in silent reading, and silent reading processes, by using the newest type of eye-camera (the oculometer, which is attached to a computer system), and which is able to locate a visual fixation of a given length of time, at a precise point in the text.

Much more research needs to be done in the area of memory span and the processing of written language by young children.

Teaching Reading

The findings of this study emphasize the importance of the language component in teaching reading—that is, especially the syntax of the language, rather than the recoding aspect of reading, which stresses word recognition skills. This study has shown that children who scored low on silent reading comprehension tended to combine syntax (or syntactic constituents) and timing cues (pauses), which were inappropriate to each other. In addition, the data also indicated that adequate comprehension in silent reading was possible without thorough word recognition skills. These data suggest strongly what Goodman (1968) and Weber (1968) have already proposed—that more emphasis should be placed on the syntactic structure of the reading text, rather than on words as such.

The investigator also feels that although oral reading may be used as a diagnostic tool, nevertheless, it may not be an accurate indication of how well a child is able to comprehend silent reading material. Since comprehension is the ultimate goal of reading, then oral reading does not appear to have any merits in attempting to teach or to test comprehension.

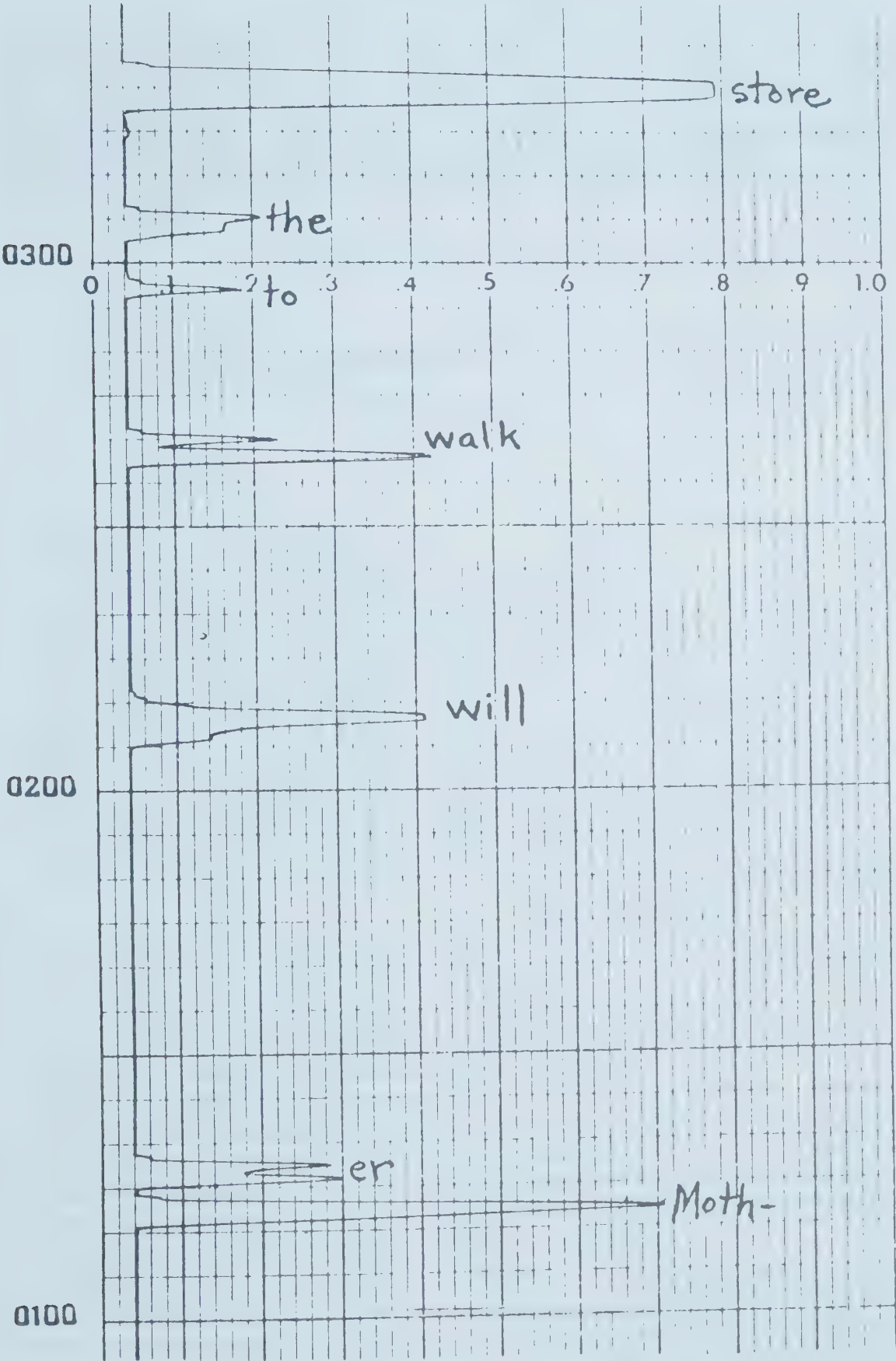
Finally, the children who scored lowest on comprehension in this study seemed to possess the best rote memories. It may be that these children tend to rely too specifically on this type of learning. Teachers should be aware of this, and test comprehension of reading material by requiring responses that do not utilize rote memory.

References

- Bond, Z. S. Units of speech perception. In *Working papers in linguistics*, No. 9. Computer and Information Science Research Centre, Ohio State University, Columbus, Ohio, 1971, 1-112.
- Brown, E. R. Lexical and syntactic predictors of the distribution of pause time in reading, (Doctoral dissertation, Syracuse University.) Ann Arbor, Mich.: University Microfilms, 1971, No. 71-18468.
- Chomsky, N., & Halle, M. *The sound pattern of English*. New York: Harper & Row, 1968.

- Cromer, W. The difference model: A new explanation for some reading difficulties. *Journal of Educational Psychology*, 1970, 61 (6), 471-484.
- Dale, H. C. A. Retroactive interference in short-term memory. *Nature*, 1964, 203, 1408.
- Fodor, J. A., & Bever, T. G. The psychological reality of linguistic segments. *Journal of Verbal Learning and Verbal Behavior*, 1965, 4, 414-420.
- Garret, M., Fodor, J. A., & Bever, T. The active use of grammar in speech perception. *Perception and Psychophysics*, 1966, 1, 30-32.
- Goldman-Eisler, F. *Psycholinguistics—Experiments in spontaneous speech*. New York: Academic Press, 1968.
- Goodman, K. S. (Ed.) *The psycholinguistic nature of the reading process*. Detroit: Wayne State University Press, 1968.
- Goodman, K. S., & Niles, O. S. *Reading process and program*. Urbana, Ill.: N. C. T. E. Publication, 1970.
- Kolers, P. A. Three stages of reading. In H. Levin & J. Williams (Eds.), *Basic Studies on reading*. New York: Basic Books, 1970.
- Latham, R. O. M. The relationship of patterns of written language segmentation to reading comprehension. A dissertation proposal submitted to the Department of Elementary Education, University of Alberta, 1971.
- Latham, R. O. M. Personal communications, "On the syntactic constituent," January 7, 1972; March 17, 1972; July 11, 1972.
- McCracken, R. A. The informal reading inventory as a means of improving instruction. *Perspectives in reading, the evaluation of children's reading achievement*, 1967, 8, 85.
- Martin, J. E., Kolodziej, B., & Genay, J. Segmentation of sentences into phonological phrases as a function of constituent length. *Journal of Verbal Learning and Verbal Behavior*, 1971, 10 (3), 226-233.
- Miller, G. A., & Isard, S. Some perceptual consequences of linguistic rules. *Journal of Verbal Learning and Verbal Behavior*, 1963, 2, 217-228.
- Neisser, U. *Cognitive psychology*. New York: Appleton-Crofts, 1967.
- Neisser, U. The role of rhythm in active verbal memory serial intrusions. *American Journal of Psychology*, 1969, 82 (4), 540-546.
- North, A. J., & Jenkins, L. B. Reading speed and comprehension as a function of typography. *Journal of Applied Psychology*, 1951, 35, 225-228.
- O'Connell, D. C., Turner, E. A., & Onuska, A. Intonation, grammatical structure, and contextural association of immediate recall. *Journal of Verbal Learning and Verbal Behavior*, 1968, 7, 110-116.
- Osgood, C. E., & Sebeok, J. A. (Eds.) *Psycholinguistics*. *Journal of Abnormal Psychology*, 1954, 49 (4, Pt. 2), 8-49.
- Pival, J. G. Stress, pitch and juncture: Tools in the diagnosis and treatment of reading ills. In E. Dechant (Ed.), *Detection and correction of reading difficulties*. New York: Appleton-Century-Crofts, 1971.
- Smith, F. *Understanding reading: A psychological analysis of reading and learning to read*. New York: Holt, Rinehart & Winston, 1971.
- VanUden, A. M. J. *A world of language for deaf children. Part 1: Basic principles*. Rotterdam: Rotterdam University Press, 1970.
- Weber, R. M. The study of oral reading errors: A survey of the literature. *Reading Research Quarterly*, 1968-69, 4, 96-119.
- Wickelgren, W. A. Acoustic similarity and retroactive interference in short-term memory. *Journal of Verbal Learning and Verbal Behavior*, 1965, 4, 53-61.
- Wickelgren, W. A. Context-sensitive coding, associative memory, and serial order in speech behavior. *Psychological Review*, 1969, 76, 1-15.
- Wilkes, A. L., & Kennedy, R. A. The relative accessibility of list items within different pause-defined groups. *Journal of Verbal Learning and Verbal Behavior*, 1970, 9, 197-201.

APPENDIX A

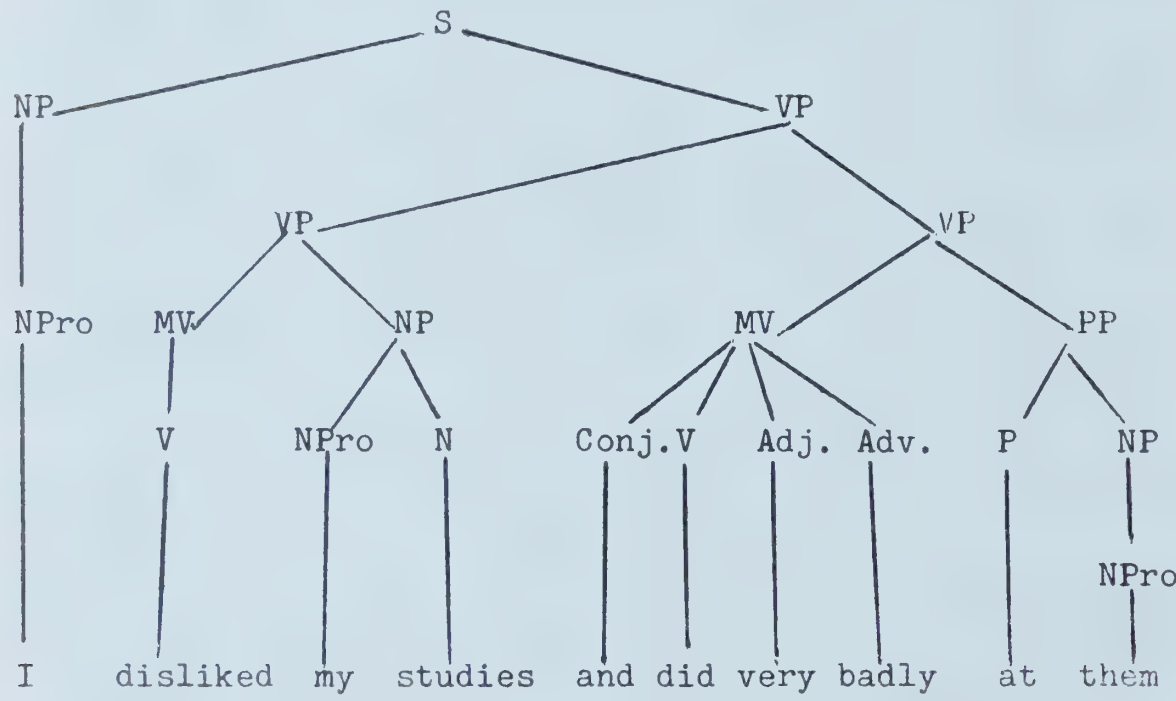


GRADE TWO BELOW-AVERAGE READER

APPENDIX B
SYNTACTIC CONSTITUENT

The term “Syntactic Constituent” will refer to the lowest major constituent as defined by Latham (1971). It may be formed from all labelled nodes in a surface structure tree. These constituents are found by locating those nodes which are immediately above the lexical nodes and deciding whether or not there are sister-nodes to the lexical nodes. If there are no sister-nodes to a specific lexical node, then the node immediately dominating that lexical node is a lowest major constituent. If there are sister nodes to a specific lexical node, then the lowest major constituent associated with that node is the node which immediately dominates all sisters of the lexical node in question.

Example: I disliked my studies and did very badly at them.



Lexical nodes: I dislike my studies and did very badly at them
Sister nodes: (1) my studies
 (2) and did very badly
The lowest major constituent associated with the sister-nodes is the NP node for (1) and the MV node for (2)

G. BROADLEY

and

KATHLEEN M. BROADLEY

County of Newell

Rural Standardisation of the Burt-Vernon Graded Word Reading Test

A rural standardisation conducted in the County of Newell, 120 miles from Calgary, confirmed the generalisability of the Burt-Vernon Graded Word Reading Test. Norms closely paralleled those established by Vernon in Calgary in 1973. A significant difference was found among younger children, however, indicating that the Calgary norms may be pro-rated too far downwards. It would appear also that with the use of a phonetically based programme, the rate of learning sight words accelerates rapidly once particular skills have been mastered. An hypothesis that girls taught by women teachers would perform significantly better than boys was not upheld. (Guy Broadley and Kathleen Broadley are with the Education Department of the County of Newell in Brooks, Alberta.)

This study was a replication of Vernon's re-standardisation of the Burt-Vernon Graded Word Reading Test (Vernon, 1973). It was conducted in the County of Newell No. 4 where the total school population is in the vicinity of 1,500 pupils. There are eight public schools in the County which covers an area of 2,350 square miles. The purpose of the replication was to compare rural Albertan reading attainments with those of Calgary children, and to provide local norms on a graded word reading test. Although other word reading tests are available, the Calgary restandardisation was being used in the County. It was felt that the norms of this test would be more accurate in an area the central point of which is only 120 miles from Calgary.

The test itself is widely used to establish word reading attainment, and is useful in diagnosing the types of errors that are involved in the child's decoding, even though its prime purpose is not diagnostic. The test is easily and quickly administered and the results, coupled with the examiner's clinical observations, often provide direction for further diagnosis of reading difficulty.

When Vernon restandardised the test in 1973, he found that the Calgary children of that year learned later than the Burt sample of 50 years before, but that the rate of learning, once begun, was faster. This was apparent in the findings until the age of 12½ was reached. At that age convergence took place. Factors such as later school entry, less emphasis on oral reading, and higher socioeconomic status were mentioned by Vernon as possible contributors to the differences found. A change in the frequencies of words recognised by the Calgary sample was reported, and a revised order of the test words was recommended for local use.

Due to the relatively standard form of curriculum and instruction in the rural and urban areas of Alberta, the authors hypothesised that overall, little difference would be found between the Calgary norms and those that would be developed from the rural sample. However, it was expected that after initial contact with formal reading instruction, when the child learned to read a few words and understand the simpler mechanics of decoding, a rapid increase in reading development would take place and that this would be reflected in the scores. It was hypothesised that Vernon's norms, through being downward pro-rated, would not reflect this phenomenon.

Vernon (1973) suggested that exposure to media may be a factor that has contributed to the establishment of the order of knowledge of words. As rural children are now exposed to much the same media as city children, it was hypothesised that the frequency of recognition of individual words in the test would be very similar to the Calgary findings.

There is already considerable evidence suggesting that boys lag behind girls in reading development, at least until the end of elementary school. Kagan (1964) suggested that the discrimination of observed sex-role activities contributed much to the difference between the reading attainments of boys and girls. Preston (1962) noted that the German boys of that time scored better in reading than the girls and that German elementary teachers were predominantly male. As all teachers but one of grades four and below within the County sample were female, it was hypothesised that girls would attain better than boys up to age ten.

Methodology

Sampling

The total population, grades one to nine, in the eight schools of the County of Newell was ascertained from school returns for October 31, 1974. The only schools in the County that were excluded were the three Hutterite schools. It was decided by the authors that 180 children would provide a manageable sample. This was approximately one child in seven of the grade one to nine population. The schools were randomly ordered and numbered. Children in the respective grades within their own schools were alphabetized regardless of sex. A number less than eight was drawn and this decided the first member of the sample, in grade one, in the first school. Every seventh child was selected thereafter in that and the succeeding grade one classes in the sample, then on through the other grades, selecting from all the schools in order at each grade level.

If a member of the sample was absent when the testing was taking place, the child immediately alphabetically prior was taken into the sample as a sub-

stitute. If this child was also absent the one immediately after the original choice was included. In these substitutions no other criteria were taken into account.

Four students were later dropped from the sample because they were the only 15-year-old children tested. They were not regarded as representative of their age group as most 15-year-old children are in grade ten. The remaining 170 students made up the final sample. They were evenly distributed in grades one through nine with the lowest number in any grade being 17, and the largest, 22. The age range was from 6 years 0 months to 14 years 10 months. There was an equal distribution of boys and girls.

Data Collection

The data were collected in one week at the end of the first school term of the 1974-75 school year. The testing was done by the authors in the following way:

The respondent was shown the printed form with the test on it (Calgary revised order was used) and asked to read each word. Responses were recorded on a duplicate form. After five consecutive errors the respondent was encouraged to look ahead in the test for other words that might be attempted. If a response on any word was not begun after fifteen seconds the child was asked to proceed to the next word.

Thus, test administration was a little different from standard procedure, but as the child still had the opportunity to make an attempt on every word, it was felt that the slightly streamlined administration procedure employed would not affect scores and would speed up the testing.

The respondent's age, sex, grade, and score (number of correctly pronounced words) were recorded on the form. To determine the reliability of the scores, the standard errors of the mean scores of three groups containing successive age groups were found. It seemed appropriate to include at least the levels immediately above and below because given chronological ages will incorporate several reading ages (and vice versa). Standard errors thus arrived at were 2.6 words (6-8 years), 2.2 words (9-11 years), and 1.7 words (12-14 years). The 5% confidence limits ranged from five words at the lower age levels to three words among the older children. The authors have integrated these standard errors into a norm structure that has converted the decimals of years into months, making comparisons with other test results a little simpler. This is available on request from the authors.

Norming

A scattergram was drawn showing age in years and months on the x axis, and score on the test on the y axis. The median age of children who obtained each score was computed. These median ages were joined on the scattergram and the curve smoothed. This is the same method of norming as used by Vernon in the 1973 Calgary restandardisation. A smoothed curve through the median ages of each group of ten scores (0-9, 10-19, etc.) proved to be identical.

Discussion of Results

The hypothesis that little difference would be found between the norms of the urban and rural samples was upheld. The two sets of scores were compared at the various median ages (reading ages). Only below median age 7½ were the

TABLE 1
A COMPARISON OF THE URBAN (VERNON, CALGARY) AND
RURAL (COUNTY OF NEWELL) NORMS OBTAINED ON THE BURT-VERNON
GRADED WORD READING TEST

Calgary 1973		County 1974		Calgary 1973		County 1974		Calgary 1973		County 1974	
R.A. ^a	Score ^b	R.A.		R.A.	Score	R.A.		R.A.	Score	R.A.	
6.3	1	6.7		7.9	37	7.8		10.8	73	10.6	
6.3	2	6.7		8.0	38	7.8		10.9	74	10.7	
6.3	3	6.7		8.1	39	7.9		11.0	75	10.8	
6.3	4	6.7		8.1	40	8.0		11.0	76	10.9	
6.3	5	6.7		8.2	41	8.1		11.1	77	11.0	
6.4	6	6.8		8.3	42	8.2		11.2	78	11.2	
6.4	7	6.8		8.3	43	8.3		11.3	79	11.3	
6.5	8	6.8		8.4	44	8.3		11.4	80	11.5	
6.5	9	6.9		8.5	45	8.4		11.5	81	11.6	
6.5	10	6.9		8.6	46	8.4		11.6	82	11.8	
6.6	11	6.9		8.7	47	8.5		11.7	83	12.0	
6.6	12	7.0		8.7	48	8.5		11.8	84	12.1	
6.6	13	7.0		8.8	49	8.6		11.9	85	12.3	
6.7	14	7.0		8.9	50	8.7		12.0	86	12.5	
6.7	15	7.1		9.0	51	8.8		12.1	87	12.6	
6.7	16	7.1		9.1	52	8.9		12.2	88	12.7	
6.8	17	7.1		9.2	53	8.9		12.3	89	12.8	
6.8	18	7.2		9.3	54	9.0		12.4	90	13.0	
6.9	19	7.2		9.3	55	9.1		12.5	91	13.1	
6.9	20	7.2		9.4	56	9.1		12.6	92	13.2	
7.0	21	7.2		9.5	57	9.2		12.7	93	13.3	
7.0	22	7.2		9.6	58	9.3		12.8	94	13.4	
7.1	23	7.2		9.7	59	9.4		13.0	95	13.4	
7.1	24	7.3		9.8	60	9.5		13.3	96	13.5	
7.2	25	7.3		9.8	61	9.6		13.5	97	13.6	
7.2	26	7.3		9.9	62	9.7		13.7	98	13.7	
7.3	27	7.3		10.0	63	9.8		14.0	99	13.8	
7.3	28	7.4		10.1	64	9.8		14.4	100	13.9	
7.4	29	7.4		10.2	65	9.9		14.7	101	14.0	
7.5	30	7.5		10.3	66	10.0		14.9	102	14.3	
7.5	31	7.5		10.3	67	10.1		15.2	103	14.6	
7.6	32	7.5		10.4	68	10.2		15.5	104	14.9	
7.7	33	7.6		10.5	69	10.2		15.8	105		
7.7	34	7.6		10.6	70	10.3		16.2	106		
7.8	35	7.7		10.6	71	10.4		16.5	107		
7.8	36	7.7		10.7	72	10.5		16.8	108		
									109		
									110		

^a Reading Age

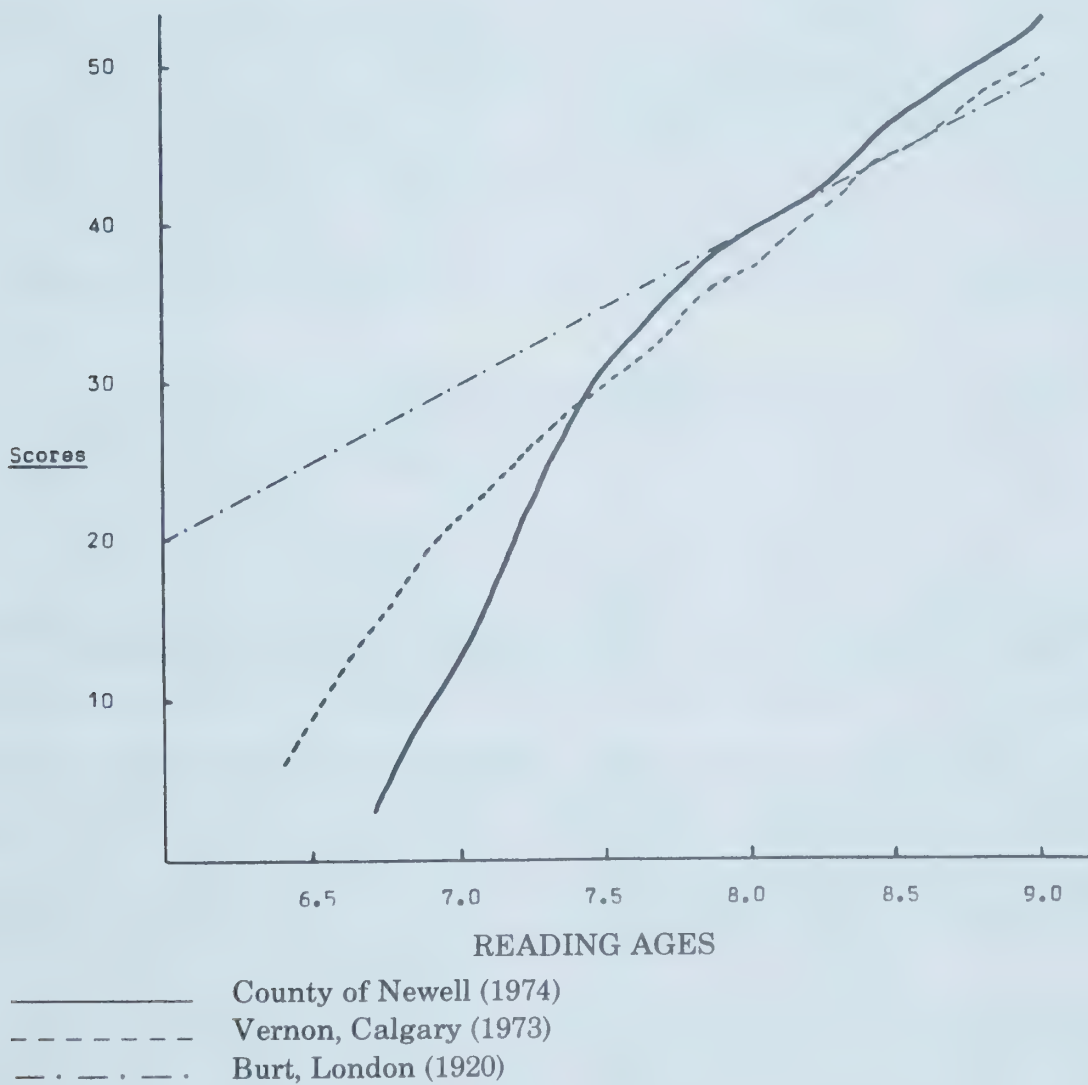
^b The score is the number of words read correctly on the test.

scores obtained markedly different. Here, the mean scores at each median age were compared from median age 6.7, the lowest in the rural sample, through to 7.5. The difference between the means of the two sets of scores at this level was significant ($t = 3.76, 16df, p < 0.01$). This lends support to the second hypothesis that Vernon's downward pro-rated norms, necessitated by later sampling, would not reflect the actual learning curve. This curve appears to demonstrate late but rapid development in the acquisition of a sight vocabulary (see Figure I). The children in the County sample did not begin to recognise words on the form until, on average, half way through year 6.

The emphasis on the phonetic structures in the grade one reading programme develops skills that the children can then apply to many words. The authors contend that the rapid increase develops in much the same way as the language explosion in very young children. Language development of

young children expands suddenly with the use of pivot words. A similar generalisation is applied by children who at first recognise a few words by sight, but then quickly expand their sight vocabulary with the identification and use of simple phonetic elements. Using this learning technique, it would seem that the students in the County sample do not begin word recognition until much later than Burt's London children of 1920, but then learn much faster, an observation made also by Vernon of his Calgary sample in 1973. Further investigation by the authors of the rate of learning sight words at this level is presently being undertaken.

FIGURE I
GRAPH COMPARING THE NORMS OBTAINED FROM THE LONDON (1920),
CALGARY (1973) AND COUNTY OF NEWELL (1974) SAMPLES,
READING AGES 6.0-9.0



The frequency of recognition of individual words was similar to that found by Vernon in 1973. Since the movement of one word is able to alter the position of several others, changes were expected. But, even with no allowances made for any change, among the first twenty words where changes seemed greatest, they were not at all significant. Two later words, LUNCHEON and TONGUE, moved 17 places further down the list. PHILOSOPHER, BINOCULAR, and

UNIQUE, appeared to be recognised more easily by the County sample, each moving 12 words or more up the list. As children are encouraged to at least look at all the words, the word order serves only to facilitate test administration.

When a comparison of boys' and girls' scores was made at the lower level of the test it was found that 10 of the 14 boys in the sample younger than age 7 years 6 months read less than ten words. Of the 12 girls in the same age group only four read less than ten words. Of the total sample, 15 out of the 85 boys read less than 30 words but only 8 out of the 85 girls read less than 30 words. Thirty words is equivalent to reading age 7.5 years. However, the difference in attainments between the boys and girls up to age 11 years was not significant. Nor was the difference up to age 7.5 years significant beyond the 0.2 level. Thus the results do not therefore establish a positive trend, but neither do they completely assuage suspicions.

Aside from the difference found at the younger age levels, the rural sample closely paralleled the urban one of 1973 in attainment on the test and in frequency of recognition of words. While no pretence is made as to the validity of Vernon's norms of 1973 outside of Alberta, there certainly appears to be a likelihood that they will be little different within Albertan school systems where children speak English, start school at six, watch television, and experience similar reading programmes. If it could be assumed that these four variables are as important as they appear, then the test's norms may have even wider validity.

The co-operation of school staffs in the County of Newell #4 is acknowledged.

References

- Asher, S. R., & Markell, R. A. Sex differences in comprehension of high- and low-interest reading materials. *Journal of Educational Psychology*, 1974, 66, 680-687.
- Braine, M. D. S. The ontogeny of English phrase structure: The first phase. *Language*, 1963, 39, 1-13.
- Kagan, J. The child's sex role classification of school objects. *Child Development*, 1964, 35, 1051-1056.
- Preston, R. C. Reading achievement of German and American children. *School and Society*, 1962, 90, 350-354.
- Vernon, P. E. A restandardisation of the Burt-Vernon Graded Word Reading Test. *Western Psychologist*, 1973, 4, 72-77.

L. E. JURY

Central York Schools

D. J. WILLOWER

and

W. J. DeLACY

The Pennsylvania State University

Teacher Self-Actualization and Pupil Control Ideology

The prediction that level of teacher self-actualization would be directly related to humanism in teacher pupil control ideology was supported by data gathered from a sample of 272 public school teachers. Expected differences in the pupil control ideology of elementary and secondary teachers were found, but level of self-actualization was not significantly different for teachers grouped dichotomously according to sex, grade levels taught, or experience. (Dr. Jury is assistant superintendent, Central York Schools, York, Pennsylvania; Dr. Willower and Dr. DeLacy are professors of education, Division of Education Policy Studies, The Pennsylvania State University.)

The concept of self-actualization, discussed earlier by other writers (cf. Goldstein, 1940), but developed chiefly by Maslow (1957, 1964, 1968, 1970, 1971) has been the object of considerable theoretical attention. However, empirical inquiries have been less common. This article reports on an investigation of the relationship between public school teachers' levels of self-actualization and their orientations toward pupil control.

We were interested in probing the relationship between teacher self-actualization and attitudes toward pupil control or discipline in part because control represents an aspect of teachers' work that earlier studies show to be of cardinal importance in school organizations. Concern with pupil control has been characterized as a salient feature of life in schools, affecting teachers' relationships with peers and other colleagues such as counselors and administrators (Willower & Jones, 1967; Willower, 1971). Further, good control is often equated with good teaching (Gordon, 1957), and it has been found that more rigid views regarding discipline are one consequence of teacher socialization (Hoy, 1968). Given social system

pressures for relatively unyielding attitudes on pupil control, the question of whether teacher level of self-actualization would be a predictor of pupil ideology seemed open enough to warrant investigation.

Conceptual Framework

Compared with a person lacking self-actualization, the self-actualizing individual has been conceptualized as one who uses his or her talents and capabilities more fully, functions relatively autonomously, tends to live in the present, and has a more benevolent outlook on human nature and on life (Maslow, 1970). The Personal Orientation Inventory (POI) developed by Shostrom (1968) serves as an operational definition of self-actualization.

Pupil control ideology has been conceptualized on a humanistic-custodial continuum. A humanistic ideology emphasizes an accepting, trustful view of pupils and optimism concerning their ability to be self-disciplined and responsible. A custodial orientation stresses the maintenance of order, distrust of pupils, and a moralistic approach to pupil control. The Pupil Control Ideology (PCI) Form (Willower, Eidell, & Hoy, 1967) furnishes an operational definition for this variable.

To examine the relationship between teacher self-actualization and pupil control ideology, the following hypothesis was proposed: level of self-actualization will be directly related to humanism in pupil control ideology.

The rationale for this hypothesis was relatively straightforward. In contrast with the less self-actualizing person, the self-actualizing individual is characterized by satisfied lower order needs and perceptions of a relatively benign environment. Hence, the self-actualizing teacher is not likely to see students as part of a threatening environment and will not view strict control of students as essential.

Further, the perceived norms of the teacher group may be especially important for the less self-actualized teacher. It has been reported that teachers tend to believe that other teachers hold comparatively custodial views on pupil control (Packard & Willower, 1972). Thus, for the non-self-actualizing teacher, the colleague group could be seen as a source of pressures for custodial orientations toward students.

Procedures

Instruments

The Personal Orientation Inventory (Shostrom, 1968) was developed as a result of a "need for a diagnostic instrument which gives . . . a measure of current level of positive health or self-actualization" (Shostrom, 1964, p. 207). It consists of 150 two-choice items selected from observed value judgments of clinically healthy and clinically troubled persons, and from theoretical formulations of writers in psychology. Although the POI has several scales and subscales, a single combined score provides an overall measure of self-actualization (Shostrom, 1968; Damm, 1972).

Shostrom (1968) reported test-retest reliabilities of .91 and .93 for the POI; claims for its validity were based upon the instrument's ability to discriminate between individuals judged by clinical psychologists to be relatively self-actualized and non-self-actualized. The highest score

theoretically possible on the POI is 150; the higher the score the more self-actualizing the respondent.

The Pupil Control Ideology Form measures educators' orientations toward discipline on a continuum having ideal-type poles represented by humanism and custodialism. The instrument consists of 20 items on a 5-point scale with response categories from "strongly agree" to "strongly disagree." The scoring range is 20-100; the higher the score, the more custodial the ideology. Split-half reliabilities ranged from .91 to .95; the instrument was shown to differentiate between teachers and schools judged to be custodial or humanistic (Willower, Eidell, & Hoy, 1967).

Sample

The instruments were administered to all of the teachers in a single school district in a northeastern state in the fall of 1972. The selected district combined suburban and rural features but bordered a middle-sized city. It had seven elementary schools, a junior and a senior high school. Faculty salary, years of teaching experience, and educational level approximated state averages. The faculty of the district was composed of 137 elementary and 141 secondary teachers.

Findings

Usable returns were secured from a total of 272 teachers, 134 at the elementary level and 138 at the secondary level. The high rate of return, 98 percent, probably resulted because the study was endorsed by both the teachers' association and the school superintendent, and the forms were completed during faculty meetings.

The main hypothesis was tested for the overall sample and separately for elementary and secondary teachers. For these tests, the Pearson product-moment correlation was the statistic employed. These and all other calculations were performed using programs of the Computation Center of The Pennsylvania State University.

Relevant data appear in Table 1. Negative correlation coefficients are a function of the scoring of the instruments. Each of the three tests of the hypothesis produced a significant association between self-actualization and pupil control ideology. Hence, the hypothesis that teacher level of self-

TABLE 1
TEACHER SELF-ACTUALIZATION AND PUPIL CONTROL IDEOLOGY

Teachers	N	SA		PCI		r
		$\bar{X}$	S.D.	$\bar{X}$	S.D.	
All	272	96.8	13.1	54.3	10.6	-.39**
Elem.	134	96.6	13.4	50.1	9.7	-.51**
Sec.	138	96.9	12.8	58.3	9.6	-.34*

* $p < .01$

** $p < .001$

actualization will be directly related to humanism in pupil control ideology cannot be rejected.

Because self-actualization (SA) has a developmental aspect and is commonly conceived as a stage in life when a mature personality has been formed, a partial correlation also was computed between SA and PCI for the overall sample, controlling for the effect of age. The partial correlation was $-.37$, significant beyond the $.001$ level. The correlation coefficient for SA and age was $-.16$, quite low but significant at the $.05$ level. This indicates a slight negative association between SA and age for the present sample of teachers.

Other demographic characteristics on which data were gathered included teacher sex, grade level taught (elementary vs. secondary), and experience (more than 5 years vs. 5 years or less). When *t*-tests were computed to examine mean differences on SA for teachers grouped according to these dichotomous variables, none were found to be significant.

On the other hand, significant differences in PCI were found on all three variables. Elementary teachers were more humanistic than their secondary school counterparts; females were more humanistic than males; and less experienced teachers were more humanistic than more experienced ones. See Jury (1973) for details. These findings are consistent with the results of previous investigations (Willower, Eidell, & Hoy, 1973).

Discussion

As predicted, teacher level of self-actualization was significantly associated with pupil control ideology: the higher the level of self-actualization, the more humanistic the ideology. The association between the two variables was strongest for elementary school teachers. That is not surprising since, in elementary schools, norms for pupil control in the teacher collectivity lay less stress on custodial views and behavior than those prevalent in secondary schools.

In this connection, it might be conjectured that more self-actualizing teachers would have fewer peer conflicts and greater work satisfaction in the elementary school than in the secondary school, and hence might be found in greater numbers in elementary schools. Yet self-actualization mean scores did not differ significantly across various teacher groupings. If only those teachers who scored very high, 112 and above, on self-actualization are considered, the distribution is quite even between elementary and secondary levels. Of these 35 self-actualizing teachers, 17 were in elementary schools and 18 were in secondary schools. The elementary school apparently does not attract and retain teachers who have higher levels of self-actualization than those in secondary school.

There have been a substantial number of inquiries on pupil control. Some of them probe relationships that stem from the character of the organizational setting and the teacher group. In general, it seems that social system factors create pressures in custodial rather than in humanistic directions.

Other studies deal with variables that are more individual and psychological in orientation. Thus, there is evidence that, contrasted with more custodial colleagues, teachers who are humanistic in pupil control ideology

tend to be less dogmatic (Willower, Eidell, & Hoy, 1967), exhibit less status obeisance or deference toward authority (Helsel, 1971a), hold emergent rather than traditional values (Helsel, 1971b), have a greater sense of power over school matters (Zelei, 1971), and score higher on measures of creativity (Halpin & Goldenberg, 1973). All of these investigations, except the last, which examined a sample of college students who planned to become teachers, were based upon relatively large samples of teachers.

The portrait of the more humanistic teacher that emerges is quite consistent with the findings of the present research. That research, which adds self-actualization to the overall picture, fleshes out a portrait composed of features that are likely to be seen as desirable. Who wants to be counted against open-mindedness, creativity, and self-actualization?

However, some basic issues remain. One deals with the relation of the teacher and the social setting. Here the organizational character of the public school is critical. Schools are characterized by vague and diverse goals, ambiguous criteria of success, lack of a widely accepted technology of teaching, a mandated relationship with clients who may or may not be willing participants, high population density, stimulus overload especially for teachers, a host of logistical problems, and political vulnerability. In such circumstances, schools like virtually all organizations, maintain structures that serve stabilizing functions. Routines, rules, and the universalistic processing of clientele are illustrative of such structures (see Willower, 1971). All this leads, in the school, to an opposition of the adult and student groups. The adults speak for the organization; the students variously resist, conform, or adapt. Casting our portrait of the more humanistic teacher against these considerations, it should be clear that constraints abound and the possibility of unanticipated negative consequences is substantial.

However, the association between pupil control ideology and pupil control behavior is problematic. This is a second basic issue. Other work shows a significant but moderate correlation between teacher pupil control ideology and behavior (Helsel & Willower, 1974). As might be expected in the social setting discussed, the correspondence is far from perfect. The results of an earlier study (Rexford, Willower, & Lynch, 1972) that showed some differences in classroom behavior between teachers who were quite humanistic and those who were quite custodial in pupil control ideology is perhaps typical of what can be expected. Using Flanders' observation techniques, it was found that the humanistic teachers were more indirect in classroom verbal behavior but did not differ significantly from the custodial teachers in percentage of teacher talk in the classroom. The school's organizational arrangements appear to afford what could be called a circumscribed leeway for teacher behavior.

A final issue concerns the consequences of teacher pupil control ideology and behavior for pupils. Empirical work is limited. Hoy (1973) reported that, in general, the more custodial the pupil control ideology of the school faculty, the more alienated the students. Pritchett (1973) found that custodialism in teacher pupil control behavior was associated with negative attitudes toward school on the part of students.

Such investigations suggest the viability of humanistic approaches to students. However, organizational and social system factors like those dis-

cussed create pressures that can mute or distort humanistic intentions and efforts. It is important to examine variables like self-actualization that can be employed to characterize teachers and that can be related to desirable teacher attitudes and behavior. Yet it is essential to see the whole within the school's social context.

References

- Damm, V. J. Overall measures of self-actualization derived from the personal orientation inventory: A replication and refinement study. *Educational and Psychological Measurement*, 1972, 32, 485-489.
- Goldstein, K. *Human nature in the light of psychopathology*. Cambridge: Harvard University Press, 1940.
- Gordon, C. W. *The social system of the high school*. Glencoe: Free Press, 1957.
- Halpin, G., & Goldenberg, R. Relationships between measures of creativity and pupil control ideology. Paper presented at the annual meeting of the American Educational Research Association, New Orleans, 1973 (mimeographed).
- Helsel, A. R. Status obeisance and pupil control ideology. *Journal of Educational Administration*, 1971, 9, 38-47. (a)
- Helsel, A. R. Value orientation and pupil control ideology of public school educators. *Educational Administration Quarterly*, 1971, 7, 24-33. (b).
- Helsel, A. R., & Willower, D. J. Toward definition and measurement of pupil control behavior. *Journal of Educational Administration*, 1974, 12, 114-123.
- Hoy, W. K. The influence of experience on the beginning teacher. *School Review*, 1968, 76, 312-323.
- Hoy, W. K. Dimensions of student alienation and characteristics of public high schools. *Interchange*, 1973, 3, 38-52.
- Jury, L. E. Teacher self-actualization and pupil control ideology. Unpublished doctoral dissertation. The Pennsylvania State University, 1973.
- Maslow, A. H. Self-actualization and beyond. In J. F. T. Bugental (Ed.), *Challenges of humanistic psychology*. New York: McGraw-Hill, 1957.
- Maslow, A. H. *Religions, values, and peak experiences*. Columbus: Ohio State University Press, 1964.
- Maslow, A. H. *Toward a psychology of being*. (2d ed.), New York: Van Nostrand Reinhold, 1968.
- Maslow, A. H. *Motivation and personality*. (2d ed.), New York: Harper and Row, 1970.
- Maslow, A. H. *The farther reaches of human nature*. New York: Viking Press, 1971.
- Packard, J. S., & Willower, D. J. Pluralistic ignorance and pupil control ideology. *Journal of Educational Administration*, 1972, 10, 78-87.
- Pritchett, W. The relationship between teacher pupil control behavior and student attitudes toward school. Unpublished doctoral dissertation. The Pennsylvania State University, 1973.
- Rexford, G. E., Willower, D. J., & Lynch, P. D. Teachers' pupil control ideology and classroom verbal behavior. *Journal of Experimental Education*, 1972, 40, 78-82.
- Shostrom, E. L. An inventory for the measurement of self-actualization. *Educational and Psychological Measurement*, 1964, 24, 207-218.
- Shostrom, E. L. *Manual for the personal orientation inventory: An inventory for the measurement of self-actualization*. San Diego: Educational and Industrial Testing Service, 1968.
- Willower, D. J. The teacher subculture. In L. W. Drabick (Ed.), *Interpreting education: A sociological approach*. New York: Appleton, Century, Crofts, 1971.
- Willower, D. J., Eidell, T. L., & Hoy, W. K. *The school and pupil control ideology*. University Park: Penn State Studies No. 24, 1967.

- Willower, D. J., Eidell, T. L., & Hoy, W. K. *The school and pupil control ideology* (2d ed. with an annotated bibliography of pupil control studies). University Park: Penn State Studies No. 24, 1973.
- Willower, D. J., & Jones, R. G. Control in an educational organization. In J. D. Raths & others (Eds.), *Studying teaching*. Englewood Cliffs, N. J.: Prentice-Hall, 1967.
- Zelei, R. A. Relationship between pupil control ideology and sense of power of teachers in selected public schools. Unpublished doctoral dissertation. University of Akron, 1971.

The Alberta Journal of Educational Research

Volume XXI, 1975

TABLE OF CONTENTS

- Arikado, M. S. *Status consensus as it relates to team teacher satisfaction*. No. 2, June, 104-109.
- Blackall, B. M. *The works of Jean Piaget published in English*. No. 1, March, 63-70.
- Breiter, J. C. *Reading or listening: A comparison of techniques of instruction in elementary social studies*. No. 2, June, 130-140.
- Broadley, G., and Broadley, K. M. *Rural standardisation of the Burt-Vernon graded word reading test*. No. 4, December, 289-294.
- Card, B. Y. *Community vocational centres in northwestern Alberta*. No. 3, September, 168-182.
- Clarke, S. C. T., and Coutts, H. T. *Toward teacher education in the year 2000*. No. 4, December, 221-240.
- Clifford, M. M. *Validity of expectation—A developmental function*. No. 1, March, 11-17.
- Das, J. P., Manos, J., and Kanungo, R. N. *Performance of Canadian native, black and white children on some cognitive and personality tests*. No. 3, September, 183-195.
- Eagan, Sr. R. *An investigation into the relationship of the pausing phenomena in oral reading and reading comprehension*. No. 4, December, 278-288.
- Eastwood, G. R. *Educational research: Concepts and models*. No. 2, June, 71-83.
- Forester, A. D. *Learning the language of reading—An exploratory study*. No. 1, March, 56-62.
- Foster, L., and Nixon, M. *The interview reassessed*. No. 1, March, 18-22.
- Friesen, D. *A study of a public school system, Hessen, West Germany*. No. 2, June, 84-103.
- Froese, V. *Does the Dolch do?* No. 1, March, 23-27.
- Hobbs, E. D. *Methodological problems in conservation testing with particular reference to volume conservation*. No. 4, December, 262-277.
- Jansson, L. C. *The judgment of simple deductive arguments by pre-service elementary school teachers*. No. 1, March, 1-10.
- Jury, L. E., Willower, D. J., and DeLacy, W. J. *Teacher self-actualization and pupil control ideology*. No. 4, December, 295-301.
- Kehoe, J. W. *Demonstrating the relationship between values and attitudes as a means of changing attitudes*. No. 3, September, 207-212.
- Klein, J. P. *Socratic dialogue vs. behavioural practice in the development of coping skills*. No. 4, December, 255-261.
- Maguire, T. O., Patsula, R. B., and Evanechko, P. O. *The development of word meaning discrimination in children*. No. 3, September, 154-167.
- Martin, J. *Controlled vs. natural setting—Some implications for behavioural analysis and change in classroom situations*. No. 1, March, 39-45.
- McCatty, C. A. *Patterns of learning projects among professional men*. No. 2, June, 116-129.
- Nelson, D., and Sawada, D. *Studying problem solving behavior in young children—Some methodological considerations*. No. 1, March, 28-38.
- Nixon, M., and Gue, L. *Women administrators and women teachers: A comparative study*. No. 3, September, 196-206.
- Ollila, L. O., and Chamberlain, L. A. *The effect of noise and object on acquisition of a sight vocabulary in kindergarten children*. No. 3, September, 213-219.

- O'Reilly, R. *Classroom climate and achievement in secondary school mathematics classes*. No. 4, December, 241-248.
- Poole, H. E. *A logically first stage in training creative teachers?* No. 3, September, 143-153.
- Pritchett, W., and Willower, D. J. *Student perceptions of teacher pupil control behavior and student attitudes toward high school*. No. 2, June, 110-115.
- Robitaille, D. F. *Classroom personality patterns of teachers of secondary mathematics*. No. 4, December, 249-254.
- Verma, S., and Peters, D. L. *Day care teacher practices and beliefs*. No. 1, March, 46-55.

New Zealand Journal of Educational Studies

Published twice yearly by the
New Zealand Council for Educational Research

Editor: P.J. Lawrence

Recent Articles

The Role of Grammar in a Secondary School English Curriculum
W.B. Elley, I.H. Barham, H. Lamb and M. Wylie

Educational Innovation: The Case of the New Zealand Playcentre
G. McDonald

The Educational Philosopher and Educational Research
R.P. Precians

The Development of Morphological Rules in Children with Different
Language Backgrounds
Marie M. Clay

Reduction of Disruption Behaviour in the Classroom: Group and
Individual Reinforcement Contingencies Compared
M.D. Ellery, N.M. Blampied and W.A.M. Black

Attending Behaviour and Reading
C.B.J. Harper and E.N. Graham

Social Participation and Special Class Attendance in
Mildly-Retarded Children
Rosemary F. Smart and Keri M. Wilton

The Effect of Frequent School Changes on Children's Self
Concept Development
G.A. Wagner and G.I. Feletti

Annual Subscription rate: Canada and the USA, US\$8.00
All business correspondence and subscriptions should be addressed to
The Executive Officer, NZCER, Box 3237, Wellington, New Zealand.

CONTEMPORARY EDUCATION

The NOW Journal
of Teaching and Learning

CE: Publishes articles by
teachers, professors and
school administrators

CE: Reviews current books
in the broad field of
education

CE: Helps educators with their
professional tasks

CE: Offers lively editorials

CE: Gives you regular columns
on provocative topics,
people

a. A View from the Bridge—
David Turney*

b. Mostly Miscellany—
M. Dale Baughman**

*Dean, School of Education, Indiana State University

**Professor of Secondary Education, and Editor of CONTEMPORARY EDUCATION

CONTEMPORARY EDUCATION, Reeve Hall,
Room 201-204, Indiana State University,
Terre Haute, Indiana 47809

Please send the next 4 issues of CONTEMPORARY EDUCATION
for \$7.00 (\$10.00 for foreign countries).

_____ Payment enclosed _____ Bill me

Name _____

Address _____

City _____

State _____ Zip _____

METHUEN ANNOUNCE

The British Journal of Teacher Education

Editor: PROFESSOR EDGAR STONES,
University of Liverpool

This new journal, designed to keep all concerned with teacher education abreast of world-wide developments in a rapidly changing field, examines the whole process of teacher education, the policies relating to its development, its institutional patterns, the changing content of initial preparation, in-service education and staff development. Two objectives of the editorial policy of the journal are that it should be actively interested in current issues, and that, with the help of a panel of overseas correspondents, it should reflect developments in teacher education on an international scale.

The journal appears three times a year, in January, April and October.

Subscription rates: Individuals £5.50; Institutions £6.75; North America £6.75 (U.S. \$18.00); airmail to all parts of the world £10.75 (U.S. \$29.00); single copies £2.25 (U.S. \$6.00). All rates inclusive of postage.

Subscriptions to: Methuen and Co., Subscriptions Department, North Way, Andover, Hampshire, SP10 5BE, England.

